

(19)



(11)

EP 1 801 223 A2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
27.06.2007 Bulletin 2007/26

(51) Int Cl.:
C12N 15/82 (2006.01)

(21) Application number: **07105544.6**

(22) Date of filing: **04.02.2004**

(84) Designated Contracting States:
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR
HU IE IT LI LU MC NL PT RO SE SI SK TR**

(30) Priority: **04.02.2003 EP 03075331**

(62) Document number(s) of the earlier application(s) in
accordance with Art. 76 EPC:
04707946.2 / 1 532 257

(71) Applicant: **CropDesign N.V.**
9052 Zwijnaarde-Gent (BE)

(72) Inventors:
• **Hatzfeld, Yves**
59000 Lille (FR)

• **Broekaert, Willem**
1700 Dilbeek (BE)

(74) Representative: **Mistry, Meeta**
BASF AG
Global Intellectual Property, GVX/C 006
67056 Ludwigshafen (DE)

Remarks:

This application was filed on 03 - 04 - 2007 as a
divisional application to the application mentioned
under INID code 62.

(54) **Rice promoters**

(57) The present invention provides several promoters isolated from *Oryza sativa*, which promoters are capable of driving and/or regulating the expression of an operably linked nucleic acid in a plant. The expression patterns of the promoters according to the present invention have been studied in *Oryza sativa* and some of the

promoters displayed specific activity in particular cells, tissues or organs of the plant, while others displayed constitutive expression throughout substantially the whole plant. Some promoters showed weak expression, while others were strongly active.

EP 1 801 223 A2

Description

[0001] The present invention relates to the field of plant molecular biology, more particularly to nucleic acid sequences useful for driving and/or regulating expression of an operably linked nucleic acid in plants. The isolation of these nucleic acid sequences from rice, as well as their use in driving and/or regulating expression of an operably linked nucleic acid is disclosed. The present invention therefore concerns promoters, hybrid promoters, genetic constructs, expression cassettes, transformation vectors, expression vectors, host cells and transgenic plants comprising the isolated nucleic acids according to the present invention. The present invention also concerns methods for driving and/or regulating expression of a nucleic acid and methods for the production of transgenic plants.

[0002] Gene expression is dependent on initiation of transcription, which is mediated via the transcription initiation complex. Gene expression is also dependent on regulation of transcription, which regulation determines how strong, when or where a gene is expressed. Said regulation of gene expression may be mediated via transcriptional control elements, which are generally embedded in the nucleic acid sequence 5'-flanking or upstream of the expressed gene. This upstream nucleic acid region is often referred to as a "promoter" since it promotes the binding, formation and/or activation of the transcription initiation complex and therefore is capable of driving and/or regulating expression of the 3' downstream nucleic acid sequence.

[0003] Genetic engineering of plants with the aim of obtaining a useful plant phenotype, often involves heterologous gene expression, which is generally mediated by a promoter capable of driving and/or regulating expression of an operably linked heterologous nucleic acid. The phenotype of the host plant only depends on the contribution of the heterologous nucleic acid, but also on the contribution of the specific expression pattern of the chosen promoter determining how, where and when that heterologous nucleic acid is expressed. Accordingly, the choice of promoter with a suitable expression pattern is of crucial importance for obtaining the suitable phenotype. A person skilled in the art will need to have available different promoters, to determine the optimal promoter for a particular nucleic acid. For many different host plants, this availability is rather limited and there is therefore a continuing need to provide new promoters with various expression profiles.

[0004] The nucleic acids as presented in SEQ ID NO 1 to 22 were isolated from *Oryza sativa* and have been found to be capable of driving and regulating expression of an operably linked nucleic acid; their expression patterns have also been characterized. Therefore the present invention offers a collection of hitherto unknown isolated nucleic acids, which isolated nucleic acids are useful as promoters.

[0005] Accordingly, the present invention provides an isolated promoter capable of driving and/or regulating expression, comprising:

(a) an isolated nucleic acid as given in any one of SEQ ID NO 1 to 22 or the complement of any one of SEQ ID NO 1 to 22; or

(b) an isolated nucleic acid having at least 90% sequence identity with any of the DNA sequences as given in any one of SEQ ID NO 1 to 22; or

(c) an isolated nucleic acid specifically hybridizing under stringent conditions with any of the DNA sequences as given in any one of SEQ ID NO 1 to 22; or

(d) an isolated nucleic acid as defined in any one of (a) to (c), which is interrupted by an intervening sequence; or

(e) a fragment of any of the nucleic acids as defined in (a) to (d), which fragment is capable of driving and/or regulating expression.

[0006] The term "isolated" as used herein means being removed from its original source. Preferably, the "isolated" promoter is free of sequences (such as protein encoding sequences or other sequences at the 3' end) that naturally flank the promoter in the genomic DNA of the organism from which the promoter is derived. Further preferably, the "isolated" promoter is also free of sequences that naturally flank it at the 5' end. Further preferably, the "isolated" promoter may comprise less than about 5 kb, 4 kb, 3 kb, 2 kb, 1.5 kb, 1.2 kb, 1 kb, 0.8 kb, 0.5 kb or 0.1 kb of nucleotide sequences that naturally occur with the promoter in genomic DNA from the organism of which the promoter is derived.

[0007] The present invention is not limited to the nucleic acids as presented by SEQ ID NO 1 to 22. A person skilled in the art will recognize that variants or fragments of a nucleic acid may occur, whilst maintaining the same functionality. These variants or fragments may be man made (e.g. by genetic engineering) or may even occur in nature. Therefore the present invention extends to variant nucleic acids and fragments of any of SEQ ID NO 1 to 22, which variants or fragments are useful in the methods of the present invention. Such variants and fragments include:

(a) an isolated nucleic acid as given in any one of SEQ ID NO 1 to 22 or the complement of any one of SEQ ID NO 1 to 22; or

(b) an isolated nucleic acid having at least 90% sequence identity with any of the DNA sequences as given in any one of SEQ ID NO 1 to 22; or

- (c) an isolated nucleic acid specifically hybridizing under stringent conditions with any of the DNA sequences as given in any one of SEQ ID NO 1 to 22; or
 (d) an isolated nucleic acid as defined in any one of (a) to (c), which is interrupted by an intervening sequence; or
 (e) a fragment of any of the nucleic acids as defined in (a) to (d), which fragment is capable of driving and/or regulating expression.

[0008] Suitable variants of any one of SEQ ID NO 1 to 22 encompass homologues which have in increasing order of preference at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with any one of the nucleic acids as represented in SEQ ID NO 1 to 22.

[0009] The percentage of identity may be calculated using an alignment program. Preferably a pair wise global alignment program may be used, which implements the algorithm of Needleman-Wunsch (J. Mol. Biol. 48: 443-453, 1970). This algorithm maximizes the number of matches and minimizes the number of gaps. Such programs are for example GAP, Needle (EMBOSS package), stretcher (EMBOSS package) or Align X (Vector NTI suite 5.5) and may use the standard parameters (for example gap opening penalty 15 and gap extension penalty 6.66). Alternatively, a local alignment program implementing the algorithm of Smith-Waterman (Advances in Applied Mathematics 2, 482-489 (1981)) may be used. Such programs are for example Water (EMBOSS package) or matcher (EMBOSS package). "Sequence identity" as used herein is preferably calculated over the entire length of the promoters as represented by any one of SEQ ID NO 1 to 22. The length of these promoters is presented in Table 2.

[0010] Search and identification of homologous nucleic acids, would be well within the realm of a person skilled in the art. Such methods, involve screening sequence databases with the sequences provided by the present invention, for example any one of SEQ ID NO 1 to 22, preferably in a computer readable form. Useful sequence databases, include but are not limited to Genbank (<http://www.ncbi.nlm.nih.gov/web/Genbank>), the European Molecular Biology Laboratory Nucleic acid Database (EMBL) (<http://w.ebi.ac.uk/ebi-docs/embl-db.html>) or versions thereof, or the MIPS database (<http://mips.gsf.de/>). Different search algorithms and software for the alignment and comparison of sequences are well known in the art. Such software includes, for example GAP, BESTFIT, BLAST, FASTA and TFASTA. Preferably BLAST software is used, which calculates percent sequence identity and performs a statistical analysis of the similarity between the sequences. The suite of programs referred to as BLAST programs has 5 different implementations: three designed for nucleotide sequence queries (BLASTN, BLASTX, and TBLASTX) and two designed for protein sequence queries (BLASTP and TBLASTN) (Coulson, Trends in Biotechnology: 76-80, 1994; Birren et al., Genome Analysis, 1: 543, 1997). The software for performing BLAST analysis is publicly available through the National Centre for Biotechnology Information.

[0011] The sequences of the genome of *Arabidopsis thaliana* and the genome of *Oryza sativa* are now available in public databases such as Genbank. Other genomes are currently being sequenced. Therefore, it is expected that as more sequences of the genomes of other plants become available, homologous promoters may be identifiable by sequence alignment with any one of SEQ ID NO 1 to SEQ ID NO 22. The skilled person will readily be able to find homologous promoters from other plant species, for example from other crop plants, such as maize. Homologous promoters from other crop plants are especially useful for practising the methods of the present invention in crop plants.

[0012] One example of homologues having at least 90% sequence identity with any one of SEQ ID NO 1 to 22 are allelic variants of any one of SEQ ID NO 1 to 22. Allelic variants are variants of the same gene occurring in two different individuals of the same species and usually allelic variants differ by slight sequence changes. Allelic variants may encompass Single Nucleotide Polymorphisms (SNPs) as well as Small Insertion/Deletion Polymorphisms (INDELs). The size of INDELs is usually less than 100 bp. SNPs and INDELs form the largest set of sequence variants in naturally occurring polymorphic strains of most organisms.

[0013] Homologues suitable for use in the methods according to the invention may readily be isolated from their source organism via the technique of PCR or hybridization. Their capability of driving and/or regulating expression may readily be determined, for example, by following the methods described in the Examples section by simply substituting the sequence used in the actual Example with the homologue.

[0014] Other suitable variants of any one of SEQ ID NO 1 to 22 encompassed by the present invention are nucleic acids specifically hybridising under stringent conditions to any one of the nucleic acids of SEQ ID NO 1 to 22. The term "hybridising" means annealing to substantially homologous complementary nucleotide sequences in a hybridization process. Tools in molecular biology relying on such a hybridization process include the polymerase chain reaction (PCR; and all methods based thereon), subtractive hybridisation, random primer extension, nuclease S1 mapping, primer extension, reverse transcription, cDNA synthesis, differential display of RNAs, and DNA sequence determination, Northern blotting (RNA blotting), Southern blotting (DNA blotting). The hybridisation process can also occur with one of the complementary nucleic acids immobilised to a matrix such as magnetic beads, Sepharose beads or any other resin. Tools in molecular biology relying on such a process include the isolation of poly (A+) mRNA. The hybridisation process can furthermore occur with one of the complementary nucleic acids immobilised to a solid support such as a nitro-cellulose or nylon membrane or immobilised by e.g. photolithography to, for example, a siliceous glass support (the

latter known as nucleic acid arrays or microarrays or as nucleic acid chips). Tools in molecular biology relying on such a process include RNA and DNA gel blot analysis, colony hybridisation, plaque hybridisation, *in situ* hybridisation and microarray hybridisation. In order to allow hybridisation to occur, the nucleic acid molecules are generally thermally or chemically denatured to melt a double strand into two single strands and/or to remove hairpins or other secondary structures from single stranded nucleic acids. The stringency of hybridisation is influenced by conditions such as temperature, salt concentration and hybridisation buffer composition. Conventional hybridisation conditions are described in, for example, Sambrook (2001) Molecular Cloning: a laboratory manual, 3rd Edition Cold Spring Harbor Laboratory Press, CSH, New York, but the skilled craftsman will appreciate that numerous different hybridisation conditions can be designed in function of the known or the expected homology and/or length of the nucleic acid sequence. High stringency conditions for hybridisation include high temperature and/or low sodium/salt concentration (salts include sodium as for example in NaCl and Na₃-citrate) and/or the inclusion of formamide in the hybridisation buffer and/or lowering the concentration of compounds such as SDS (sodium dodecyl sulphate detergent) in the hybridisation buffer and/or exclusion of compounds such as dextran sulphate or polyethylene glycol (promoting molecular crowding) from the hybridisation buffer. Specifically hybridising under stringent conditions means that the sequences have to be very similar. Specific hybridization under stringent conditions is preferably carried out at a temperature of 60°C followed by washes in 0.1 to 1 XSSC, 0.1XSDS, and 1X SSC, 0.1X SDS.

[0015] The invention also relates to a nucleic acid molecule of at least 15 nucleotides in length hybridizing specifically with any of the nucleic acids of the invention. The invention also relates to a nucleic acid molecule of at least 15 nucleotides in length specifically amplifying a nucleic acid of the invention by polymerase chain reaction.

[0016] Another variant of any of SEQ ID NO 1 to 22 encompassed by the present invention are nucleic acids corresponding to any one of SEQ ID NO 1 to 22 or variants thereof as described hereinabove, which are interrupted by an intervening sequence. For example, any of the nucleic acids as presented in SEQ ID NO 1 to 22 may be interrupted by an intervening sequence. With "intervening sequences" is meant any nucleic acid or nucleotide, which disrupts another sequence. Examples of intervening sequences comprise introns, nucleic acid tags, T-DNA and mobilizable nucleic acids sequences such as transposons or nucleic acids that can be mobilized via recombination. Examples of particular transposons comprise *Ac* (activator), *Ds* (Dissociation), *Spm* (suppressor-Mutator) or *En*. The introduction of introns into promoters is now widely applied. The methods according to the present invention may also be practised using a nucleic acid sequence according to any one of SEQ ID NO 1 to 22 provided with an intron. In case the intervening sequence is an intron, alternative splice variants of the nucleic acids according to the invention may arise. The term "alternative splice variant" as used herein encompasses variants of a nucleic acid sequence in which intervening introns have been excised, replaced or added. Such splice variants may be found in nature or may be manmade. Methods for making such promoters with an intron or for making the corresponding splice variants are well known in the art.

[0017] Variants interrupted by an intervening sequence, suitable for use in the methods according to the invention may readily be determined for example by following the methods described in the Examples section by simply substituting the sequence used in the actual Example with the variant.

[0018] The variant nucleic acids as described hereinabove may be found in nature (for example allelic variants or splice variants). Additionally and/or alternatively, variants of any one of SEQ ID NO 1 to 22 as described hereinabove may be manmade via techniques well known in the art involving for example mutation, substitution, insertion, deletions or derivation. The present invention also encompasses such variants, as well as their use in the methods of the present invention.

[0019] A "mutation variant" of a nucleic acid may readily be made using recombinant DNA manipulation techniques or nucleotide synthesis. Examples of such techniques include site directed mutagenesis via M13 mutagenesis, T7-Gen *in vitro* mutagenesis (USB, Cleveland, OH), QuickChange Site Directed mutagenesis (Stratagene, San Diego, CA), PCR-mediated site-directed mutagenesis or other site-directed mutagenesis protocols. Alternatively, the nucleic acid of the present invention may be randomly mutated.

[0020] A "substitutional variant" refers to those variants in which at least one residue in the nucleic acid sequence has been removed and a different residue inserted in its place. Nucleic acid substitutions are typically of single residues, but may be clustered depending upon functional constraints placed upon the nucleic acid sequence; insertions usually are of the order of about 1 to about 10 nucleic acid residues, and deletions can range from about 1 to about 20 residues.

[0021] An "insertional variant" of a nucleic acid is a variant in which one or more nucleic acid residues are introduced into a predetermined site in that nucleic acid. Insertions may comprise 5'-terminal and/or 3'-terminal fusions as well as intra-sequence insertions of single or multiple nucleotides. Generally, insertions within the nucleic acid sequence will be smaller than 5'- or 3'-terminal fusions, of the order of about 1 to 10 residues. Examples of 5'- or 3'-terminal fusions include the coding sequences of binding domains or activation domains of a transcriptional activator as used in the yeast two-hybrid system or yeast one-hybrid system, or of phage coat proteins, (histidine)₆-tag, glutathione S-transferase-tag, protein A, maltose-binding protein, dihydrofolate reductase, Tag*100 epitope, c-myc epitope, FLAG[®]-epitope, lacZ, CMP (calmodulin-binding peptide), HA epitope, protein C epitope and VSV epitope.

[0022] The term "derivative" of a nucleic acid may comprise substitutions, and/or deletions and/or additions of naturally

and non-naturally occurring nucleic acid residues compared to the natural nucleic acid. Derivatives may, for example, comprise methylated nucleotides, or artificial nucleotides.

[0023] Also encompassed within the present invention are promoters, comprising a fragment of any of the nucleic acids as presented by any one of SEQ ID NO 1 to 22 or variants thereof as described hereinabove. A "fragment" as used herein means a portion of a nucleic acid sequence. Suitable fragments useful in the methods of the present invention are functional fragments, which retain at least one of the functional parts of the promoter and hence are still capable of driving and/or regulating expression. Examples of functional fragments of a promoter include the minimal promoter, the upstream regulatory elements, or any combination thereof.

[0024] Suitable fragments may range from at least about 20 base pairs or about 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950 or 1000 base pairs, up to about the full length sequence of the invention. These base pairs are typically immediately upstream of the transcription initiation start, but alternatively may be from anywhere in the promoter sequence.

[0025] Suitable fragments useful in the methods of the present invention may be tested for their capability of driving and/or regulating expression by standard techniques well known to the skilled person, or by the following method described in the Example section.

[0026] The promoters as disclosed in any one of SEQ ID NO 1 to 22 are isolated as nucleic acids of approximately 1.2kb from the upstream region of particular rice coding sequences (CDS). These nucleic acids may include typical elements of a promoter, which are presented in Figure 1. Generally, a promoter may comprise from coding sequence to the upstream direction: (i) an 5'UTR of pre-messenger RNA, (ii) a minimal promoter comprising the transcription initiation element (INR) and more upstream a TATA box, and (iii) may contain regulatory elements that determine the specific expression pattern of the promoter.

[0027] The term "promoter" as used herein is taken in a broad context and refers to regulatory nucleic acid sequences capable of effecting (driving and/or regulating) expression of the sequences to which they are operably linked. A "promoter" encompasses transcriptional regulatory sequences derived from a classical genomic gene. Usually a promoter comprises a TATA box, which is capable of directing the transcription initiation complex to the appropriate transcription initiation start site. However, some promoters do not have a TATA box (TATA-less promoters), but are still fully functional for driving and/or regulating expression. A promoter may additionally comprise a CCAAT box sequence and additional regulatory elements (i.e. upstream activating sequences or cis-elements such as enhancers and silencers). A "promoter" may also include the transcriptional regulatory sequences of a classical prokaryotic gene, in which case it may include a -35 box sequence and/or a -10 box transcriptional regulatory sequences.

[0028] "Driving expression" as used herein means promoting the transcription of a nucleic acid.

[0029] "Regulating expression" as used herein means influencing the level, time or place of transcription of a nucleic acid. The promoters of the present invention may thus be used to increase, decrease or change in time and/or place transcription of a nucleic acid. For example, they may be used to limit the transcription to certain cell types, tissues or organs, or during a certain period of time, or in response to certain environmental conditions.

[0030] The promoter is preferably a plant-expressible promoter. The term "plant-expressible" means being capable of regulating expression in a plant, plant cell, plant tissue and/or plant organ. Accordingly, the invention encompasses an isolated nucleic acid as mentioned above, capable of regulating transcription of an operably linked nucleic acid in a plant or in one or more particular cells, tissues or organs of a plant.

[0031] The expression pattern of the promoters according to the present invention were studied in detail and it was found that many of them were tissue-specific. Accordingly, the present invention provides "tissue-specific" promoters. The term "tissue-specific" shall be taken to indicate that expression is predominantly in a particular tissue, tissue-type, organ or any other part of the organism, albeit not necessarily exclusively in said tissue, tissue-type, organ or other part. Accordingly, the invention encompasses an isolated nucleic acid as mentioned above, capable of driving and/or regulating expression (of an operably linked nucleic acid) in a tissue-specific manner. Expression may be driven and/or regulated in the seed, embryo, scutellum, aleurone, endosperm, leaves, flower, calli, meristem, shoot meristem, discriminating centre, shoot, shoot meristem and root. In grasses the shoot meristem is located in the so-called discrimination zone from where the shoot and the leaves originate.

[0032] A tissue-specific promoter is one example of a so-called "regulated promoter". These promoters are regulated by endogenous signals such as the presence of certain transcription factors, metabolites, plant hormones, or exogenous signals, such as ageing, stresses or nutritional status. These regulations may have an effect on one or more different levels such spatial specificity or temporal specificity. Encompassed within the present invention is a nucleic acid as described hereinabove, which is a "regulated promoter". Examples of regulated promoters are cell-specific promoters, tissue-specific promoters, organ-specific promoters, cell cycle-specific promoters, inducible promoters or young tissue-specific promoters.

[0033] Alternatively and/or additionally, some promoters of the present invention display a constitutive expression pattern. Accordingly, the present invention provides a promoter as described hereinabove, which is a constitutive promoter. The term "constitutive" means having no or very few spatial or temporal regulations. The term "constitutive

expression" as used herein refers to a substantially continuously expression in substantially all tissues of the organism. The skilled craftsman will understand that a "constitutive promoter" is a promoter that is active during most, but not necessarily all, phases of growth and development of the organism and throughout most, but not necessarily all, parts of an organism.

[0034] The "expression pattern" of a promoter is not only influenced by the spatial and temporal aspects, but also by the level of expression. The level of expression is determined by the so-called "strength" of a promoter. Depending on the resulting expression level, a distinction is made herein between "weak" or "strong" promoters. Generally by "weak promoter" is meant a promoter that drives expression of an operably linked nucleic acid at levels of about 1/10000 transcripts to about 1/100000 transcripts to about 1/500000 transcripts. Generally, by "strong promoter" is meant a promoter that drives expression at levels of about 1/10 transcripts, to about 1/100 or to about 1/1000 transcripts.

[0035] According to a particular embodiment, the invention provides an isolated promoter as mentioned hereinabove, which is a hybrid promoter. The term "hybrid promoter" as used herein refers to a chimeric promoter made, for example, synthetically, for example by genetic engineering. Preferred hybrid promoters according to the present invention comprise a part, preferably a functional part, of one of the promoters according to the present invention and at least another part, preferably a functional part of a promoter. The latter part, may be a part of any promoter, including any one of the promoters according to the present invention and other promoters. One example of a hybrid promoter comprises regulatory element(s) of a promoter according to the present invention combined with the minimal promoter of another promoter. Another example of a hybrid promoter is a promoter comprising additional regulatory elements to further enhance its activity and/or to alter its spatial and/or temporal expression pattern.

[0036] The present invention also provides use of a functional fragment of any one of SEQ ID NO 1 to 22 or variant thereof for changing the expression pattern of a promoter. In such methods, at least part of any of the nucleic acids according to the present invention are combined with at least one fragment of another promoter.

[0037] Further, the invention provides a genetic construct comprising:

- (a) An isolated promoter as defined hereinabove
- (b) A heterologous nucleic acid sequence operably linked to isolated promoter of (a), and optionally
- (c) A 3' transcription terminator

[0038] The term "genetic construct" as used herein means a nucleic acid made by genetic engineering.

[0039] The term "operably linked" to a promoter as used herein means that the transcription is driven and/or regulated by that promoter. A person skilled in the art will understand that being operably linked to a promoter preferably means that the promoter is positioned upstream (i.e. at the 5'-end) of the operably linked nucleic acid. The distance to the operably linked nucleic acid may be variable, as long as the promoter of the present invention is capable of driving and/or regulating the transcription of the operably linked nucleic acid. For example, between the promoter and the operably linked nucleic acid, there might be a cloning site, an adaptor, a transcription or translation enhancer.

[0040] The operably linked nucleic acid may be any coding or non-coding nucleic acid. The operably linked nucleic acid may be in the sense or in the anti-sense direction. Typically in the case of genetic engineering of host cells, the operably linked nucleic acid is to be introduced into the host cell and is intended to change the phenotype of the host cell. Alternatively, the operably linked nucleic acid is an endogenous nucleic acid from the host cell.

[0041] The term "heterologous" as used herein is intended to be "heterologous to the promoter of the present invention". A nucleic acid that is heterologous to the promoter of the present invention is not naturally occurring in the nucleic acid sequences flanking the promoter of the present invention when it is in its biological genomic environment. While the nucleic acid may be heterologous to the promoter of the present invention, it may be homologous or native or heterologous or foreign to the plant host cell. The heterologous operably linked nucleic acid may be any nucleic acid (for example encoding any protein), provided that it comprises or it is flanked by at least one nucleotide which is normally not flanking the promoter of the present invention.

[0042] The term "transcription terminator" as used in (c) refers to a DNA sequence at the end of a transcriptional unit which signals termination of transcription. Terminators are 3'-non-translated DNA sequences usually containing a polyadenylation signal, which facilitates the addition of polyadenylate sequences to the 3'-end of a primary transcript. Terminators active in and/or isolated from viruses, yeasts, moulds, bacteria, insects, birds, mammals and plants are known and have been described in literature. Examples of terminators suitable for use in the genetic constructs of the present invention include the *Agrobacterium tumefaciens* nopaline synthase (NOS) gene terminator, the *Agrobacterium tumefaciens* octopine synthase (OCS) gene terminator sequence, the Cauliflower mosaic virus (CaMV) 35S gene terminator sequence, the *Oryza sativa* ADP-glucose pyrophosphorylase terminator sequence (t3'Bt2), the *Zea mays* zein gene terminator sequence, the *rbcs-1A* gene terminator, and the *rbcs-3A* gene terminator sequences, amongst others.

[0043] The present invention also provides an expression cassette, a transformation vector or a plant expression vector comprising a genetic construct as described above.

[0044] An "expression cassette" as meant herein refers to a minimal genetic construct necessary for expression of a

nucleic acid. A typical expression cassette comprises a promoter-gene-terminator combination. An expression cassette may additionally comprise cloning sites, for example GatewayTM recombination sites or restriction enzyme recognition sites, to allow easy cloning of the operably linked nucleic acid or to allow the easy transfer of the expression cassette into a vector. An expression cassette may further comprise 5' untranslated regions, 3' untranslated regions, a selectable marker, transcription enhancers or translation enhancers.

[0045] With "transformation vector" is meant a genetic construct, which may be introduced in an organism by transformation and may be stably maintained in said organism. Some vectors may be maintained in for example *Escherichia coli*, *A. tumefaciens*, *Saccharomyces cerevisiae* or *Schizosaccharomyces pombe*, while others such as phagemids and cosmid vectors, may be maintained in bacteria and/or viruses. Transformation vectors may be multiplied in their host cell and may be isolated again therefrom to be transformed into another host cell. Vector sequences generally comprise a set of unique sites recognized by restriction enzymes, the multiple cloning site (MCS), wherein one or more non-vector sequence(s) can be inserted. Vector sequences may further comprise an origin of replication which is required for maintenance and/or replication in a specific host cell. Examples of origins of replication include, but are not limited to, the f1-ori and colE1.

[0046] "Expression vectors" form a subset of transformation vectors, which, by virtue of comprising the appropriate regulatory sequences, enable expression of the inserted non-vector sequence(s). Expression vectors have been described which are suitable for expression in bacteria (e.g. *E. coli*), fungi (e.g. *S. cerevisiae*, *S. pombe*, *Pichia pastoris*), insect cells (e.g. baculoviral expression vectors), animal cells (e.g. COS or CHO cells) and plant cells. One suitable expression vector according to the present invention is a plant expression vector, useful for the transformation of plant cells, the stable integration in the plant genome, the maintenance in the plant cell and the expression of the non-vector sequences in the plant cell.

[0047] Typically, a plant expression vector according to the present invention comprises a nucleic acid of any one of SEQ ID NO 1 to 22 or a variant thereof as described hereinabove, optionally operably linked to a second nucleic acid. Typically, a plant expressible vector according to the present invention, further comprises T-DNA regions for stable integration into the plant genome (for example the left border and the right border regions of the Ti plasmid).

[0048] The genetic constructs of the invention may further comprise a "selectable marker". As used herein, the term "selectable marker" includes any gene, which confers a phenotype to a cell in which it is expressed, to facilitate the identification and/or selection of cells that are transfected or transformed. Suitable markers may be selected from markers that confer antibiotic or herbicide resistance. Cells containing the genetic construct will thus survive antibiotics or herbicide concentrations that kill untransformed cells. Examples of selectable marker genes include genes conferring resistance to antibiotics (such as nptII encoding neomycin phosphotransferase capable of phosphorylating neomycin and kanamycin, or hpt encoding hygromycin phosphotransferase capable of phosphorylating hygromycin), to herbicides (for example bar which provides resistance to Basta; aroA or gox providing resistance against glyphosate), or genes that provide a metabolic trait (such as manA that allows plants to use mannose as sole carbon source). Visual marker genes result in the formation of colour (for example beta-glucuronidase, GUS), luminescence (such as luciferase) or fluorescence (Green Fluorescent Protein, GFP, and derivatives thereof). Further examples of suitable selectable marker genes include the ampicillin resistance (Ampr), tetracycline resistance gene (Tcr), bacterial kanamycin resistance gene (Kanr), phosphinothricin resistance gene, and the chloramphenicol acetyltransferase (CAT) gene, amongst others.

[0049] Furthermore, the present invention encompasses a host cell comprising an isolated promoter, or a genetic construct, or an expression cassette, or a transformation vector or an expression vector according to the invention as described hereinabove. In particular embodiments of the invention, the host cell is selected from bacteria, algae, fungi, yeast, plants, insect or animal host cells.

[0050] In one particular embodiment, the invention provides a transgenic plant cell comprising an isolated promoter according to the invention, or an isolated nucleic acid, or a genetic construct, or an expression cassette, or a transformation vector or an expression vector according to the invention as described hereinabove. Preferably said plant cell is a dicot plant cell or a monocot plant cell, more preferably a cell of any of the plants as mentioned herein. Preferably, in the transgenic plant cell according to the invention, the promoter or the genetic construct of the invention is stably integrated into the genome of the plant cell.

[0051] The invention also provides a method for the production of a transgenic plant, comprising:

- (a) Introducing into a plant cell an isolated promoter, for example any one of SEQ ID NO 1 to SEQ ID NO 22, or a variant or fragment thereof, or a genetic construct, or an expression cassette, or a transformation vector or an expression vector according to the present invention and as described hereinabove, and
- (b) Cultivating said plant cell under conditions promoting plant growth.

[0052] "Introducing" the above mentioned isolated promoter, or genetic construct, or expression cassette, or transformation vector or expression vector, into a host cell (e.g. plant cell) is preferably achieved by transformation. The term "transformation" as used herein encompasses the transfer of an exogenous polynucleotide into a host cell, irrespective

of the method used for transfer. In particular for plants, tissues capable of clonal propagation, whether by organogenesis or embryogenesis, are suitable to be transformed with a genetic construct of the present invention and a whole plant may be regenerated therefrom. The particular tissue chosen will vary depending on the clonal propagation systems available for, and best suited to, the particular plant species being transformed. Exemplary tissue targets include leaf disks, pollen, embryos, cotyledons, hypocotyls, megagametophytes, callus tissue, existing meristematic tissue (e.g., apical meristem, axillary buds, and root meristems), and induced meristem tissue (e.g., cotyledon meristem and hypocotyl meristem). The polynucleotide may be transiently or stably introduced into a plant cell and may be maintained non-integrated, for example, as a plasmid. Alternatively, it may be integrated into the plant genome.

[0053] Transformation of a plant species is now a fairly routine technique. Advantageously, any of several transformation methods may be used to introduce the nucleic acids of the invention into a suitable ancestor cell. Transformation methods include the use of liposomes, electroporation, chemicals that increase free DNA uptake, injection of the DNA directly into the plant, particle gun bombardment, transformation using viruses or pollen and microprojection. Methods may be selected from the calcium/polyethylene glycol method for protoplasts (Krens, F.A. et al., 1982, Nature 296, 72-74; Negrutiu I. et al., June 1987, Plant Mol. Biol. 8, 363-373); electroporation of protoplasts (Shillito R.D. et al., 1985 Bio/Technol 3, 1099-1102); microinjection into plant material (Crossway A. et al., 1986, Mol. Gen Genet 202, 179-185); DNA or RNA-coated particle bombardment (Klein T.M. et al., 1987, Nature 327, 70) infection with (non-integrative) viruses and the like. A preferred transformation method for the production of transgenic plant cells according to the present invention, is an *Agrobacterium* mediated transformation method.

[0054] Transgenic rice plants comprising any one of the promoters of the present invention are preferably produced via *Agrobacterium*-mediated transformation using any of the well-known methods for rice transformation, such as the ones described in any of the following: published European patent application EP 1198985 A1, Aldemita and Hodges (Planta, 199, 612-617, 1996); Chan et al. (Plant Mol. Biol. 22 (3) 491-506, 1993); Hiei et al. (Plant J. 6 (2) 271-282, 1994); which disclosures are incorporated by reference herein as if fully set forth. In the case of corn transformation, the preferred method is as described in either Ishida et al. (Nat. Biotechnol. 1996 Jun; 14(6): 745-50) or Frame et al. (Plant Physiol. 2002 May; 129(1): 13-22), which disclosures are incorporated by reference herein as if fully set forth.

[0055] Generally after transformation, plant cells or cell groupings are selected for the presence of one or more markers which are encoded by plant-expressible genes co-transferred with the gene of interest (which could be under the control of any of the promoters of the present invention), following which the transformed material may be cultivated under conditions promoting plant growth.

[0056] The resulting transformed plant cell may then be used to regenerate a transformed plant in a manner known to persons skilled in the art. Accordingly, the method for the production of a transgenic plant as described hereinabove, may further comprise regenerating a plant from said plant cell of (a).

[0057] The present invention further provides a plant comprising a plant cell as described hereinabove. The plants may also be able to grow, or even reach maturity including for example fruit production, seed formation, seed ripening and seed setting.

[0058] Furthermore, progeny may be produced from these seeds, which progeny may be fertile. Alternatively or additionally, the transformed and regenerated plants may also produce progeny by non-sexual propagation such as cloning, grafting. The generated transformed plants may be propagated by a variety of means, such as by clonal propagation or classical breeding techniques. For example, a first generation (or T1) transformed plant may be selfed to give homozygous second generation (or T2) transformants, and the T2 plants further propagated through classical breeding techniques.

[0059] The generated transformed organisms may take a variety of forms. For example, they may be chimeras of transformed cells and non-transformed cells; clonal transformants (e.g., all cells transformed to contain the expression cassette); grafts of transformed and untransformed tissues (e.g., in plants, a transformed rootstock grafted to an untransformed scion).

[0060] Following DNA transfer and growth of the transformed cells, putatively transformed plant cells or plants may be evaluated, for instance using Southern analysis, for the presence of the gene of interest, copy number and/or genomic organization. Alternatively or additionally, expression levels or expression patterns of the newly introduced DNA may be undertaken using northern and/or Western analysis, both techniques being well known to persons having ordinary skill in the art.

[0061] The present invention clearly extends to plants obtainable by any of the methods according to the present invention, which plants comprise any of the isolated promoters or the constructs of the present invention. The present invention clearly extends to any plant parts and propagules of such plant. The present invention extends further to encompass the progeny of a primary transformed cell, tissue, organ or whole plant that has been produced by any of the aforementioned methods, the only requirement being that progeny exhibit the same genotypic and/or phenotypic characteristic(s) as those produced in the parent by the methods according to the invention. The invention also extends to harvestable parts of a plant, such as but not limited to seeds, leaves, fruits, flowers, stem cultures, stem, rhizomes, roots, tubers, bulbs and cotton fibers.

[0062] The term "plant" or "plants" as used herein encompasses whole plants, ancestors and progeny of plants and plant parts, including seeds, shoots, stems, roots (including tubers), and plant cells, tissues and organs. The term "plant" therefore also encompasses suspension cultures, embryos, meristematic regions, callus tissue, gametophytes, sporophytes, pollen, and microspores. Plants that are particularly useful in the methods of the invention include all plants which belong to the superfamily *Viridiplantae*, in particular monocotyledonous and dicotyledonous plants including a fodder or forage legume, ornamental plant, food crop, tree, or shrub selected from the list comprising *Acacia* spp., *Acer* spp., *Actinidia* spp., *Aesculus* spp., *Agathis australis*, *Albizia amara*, *Alsophila tricolor*, *Andropogon* spp., *Arachis* spp., *Areca catechu*, *Astelia fragrans*, *Astragalus cicer*, *Baikiaea plurijuga*, *Betula* spp., *Brassica* spp., *Bruguiera gymnorhiza*, *Burkea africana*, *Butea frondosa*, *Cadaba farinosa*, *Calliandra* spp., *Camellia sinensis*, *Canna indica*, *Capsicum* spp., *Cassia* spp., *Centroema pubescens*, *Chaenomeles* spp., *Cinnamomum cassia*, *Coffea arabica*, *Colophospermum mopane*, *Coronilla varia*, *Cotoneaster serotina*, *Crataegus* spp., *Cucumis* spp., *Cupressus* spp., *Cyathea dealbata*, *Cydonia oblonga*, *Cryptomeria japonica*, *Cymbopogon* spp., *Cynthea dealbata*, *Cydonia oblonga*, *Dalbergia monetaria*, *Davallia divaricata*, *Desmodium* spp., *Dicksonia squarosa*, *Diheteropogon amplexans*, *Dioclea* spp., *Dolichos* spp., *Dorycnium rectum*, *Echinochloa pyramidalis*, *Ehretia* spp., *Eleusine coracana*, *Eragrostis* spp., *Erythrina* spp., *Eucalyptus* spp., *Euclea schimperi*, *Eulalia villosa*, *Fagopyrum* spp., *Feijoa sellowiana*, *Fragaria* spp., *Flemingia* spp., *Freyinetia banksii*, *Geranium thunbergii*, *Ginkgo biloba*, *Glycine javanica*, *Gliricidia* spp., *Gossypium hirsutum*, *Grevillea* spp., *Guibourtia coleosperma*, *Hedysarum* spp., *Hemarthria altissima*, *Heteropogon contortus*, *Hordeum vulgare*, *Hyparrhenia rufa*, *Hypericum erectum*, *Hyperthelia dissoluta*, *Indigo incarnata*, *Iris* spp., *Leptarrhena pyrolifolia*, *Lespedeza* spp., *Lettuca* spp., *Leucaena leucocephala*, *Loudetia simplex*, *Lotonus bainesii*, *Lotus* spp., *Macrotyloma axillare*, *Malus* spp., *Manihot esculenta*, *Medicago sativa*, *Metasequoia glyptostroboides*, *Musa sapientum*, *Nicotianum* spp., *Onobrychis* spp., *Ornithopus* spp., *Oryza* spp., *Peltophorum africanum*, *Pennisetum* spp., *Persea gratissima*, *Petunia* spp., *Phaseolus* spp., *Phoenix canariensis*, *Phormium cookianum*, *Photinia* spp., *Picea glauca*, *Pinus* spp., *Pisum sativum*, *Podocarpus totara*, *Pogonarthria fleckii*, *Pogonarthria squarrosa*, *Populus* spp., *Prosopis cineraria*, *Pseudotsuga menziesii*, *Pterolobium stellatum*, *Pyrus communis*, *Quercus* spp., *Rhaphiolepis umbellata*, *Rhopalostylis sapida*, *Rhus natalensis*, *Ribes grossularia*, *Ribes* spp., *Robinia pseudoacacia*, *Rosa* spp., *Rubus* spp., *Salix* spp., *Schyzachyrium sanguineum*, *Sciadopitys verticillata*, *Sequoia sempervirens*, *Sequoiadendron giganteum*, *Sorghum bicolor*, *Spinacia* spp., *Sporobolus fimbriatus*, *Stiburus alopecuroides*, *Stylosanthes humilis*, *Tadehagi* spp., *Taxodium distichum*, *Themeda triandra*, *Trifolium* spp., *Triticum* spp., *Tsuga heterophylla*, *Vaccinium* spp., *Vicia* spp., *Vitis vinifera*, *Watsonia pyramidata*, *Zantedeschia aethiopica*, *Zea mays*, amaranth, artichoke, asparagus, broccoli, brussel sprout, cabbage, canola, carrot, cauliflower, celery, collard greens, flax, kale, lentil, oilseed rape, okra, onion, potato, rice, soybean, straw, sugarbeet, sugar cane, sunflower, tomato, squash, and tea, trees and algae amongst others. According to a preferred feature of the present invention, the plant is a crop plant such as soybean, sunflower, canola, alfalfa, rapeseed, cotton, tomato, potato, tobacco, squash, papaya, poplar, leguminosa, flax, lupinus or sorghum. According to another preferred embodiment of the present invention the plant is a monocotyledonous plant, such as sugarcane, further preferable a cereal such as rice, maize, wheat, barley, millet, rye or oats.

[0063] The invention further provides a method for driving and/or regulating expression of a nucleic acid in a plant or plant cell, comprising:

- a) Operably linking a nucleic acid to an isolated nucleic acid according to the invention as described hereinabove, such as to any one of SEQ ID NO 1 to 22 or a variant or fragment thereof, and
- b) Introducing the resultant genetic construct into a plant or plant cell.

[0064] Preferably the operably linked nucleic acid of (a) is heterologous to the nucleic acids according to the present invention.

[0065] This method may further comprise cultivating the transformed plant or plant cell under conditions promoting growth, promoting regeneration and/or promoting maturation.

[0066] Furthermore, the expression of the operably linked nucleic acid may be driven and/or regulated in particular cells, tissues or organs of a plant. Accordingly, the invention provides a method as described above, wherein the expression is constitutive expression or tissue-specific expression. For these embodiments, reference is made to the example section where the specific expression patterns of the promoters according to the invention are described and where different types of tissue-specific expression are detailed.

[0067] The present invention further encompasses the use of an isolated nucleic acid as defined hereinabove to drive and/or regulate expression of an operably linked nucleic acid.

[0068] The person skilled in the art will recognize that provision of sequences SEQ ID NO 1 to 22, readily makes available the tools to isolate related promoters, which may have substantial sequence identity to any of SEQ ID NO 1 to 22. Additionally, provision of sequences SEQ ID NO 23 to 44 (CDS corresponding to the promoters of the present invention, see Table 1), readily makes available the tools to isolate related promoters, of which the related CDSs may have substantial sequence identity to any of SEQ ID NO 23 to 44. Therefore the present invention also encompasses

a method for isolating nucleic acids, capable of driving and/or regulating expression of an operably linked nucleic acid, comprising screening a nucleic acid sequence database to find homologues of any of the sequences represented by SEQ ID NO 1 to 22 or SEQ ID NO 23 to 44. Subsequently these homologues are used to screen a library with genomic DNA, which library is for example prepared from the organism of origin of the above mentioned homologue. The screening procedure may for example involve hybridization. Subsequently, the genomic DNA that matches the homologue, is analysed to identify the transcription initiation site and the translation initiation site of the gene corresponding to the homologue. Finally, specific primers are designed for amplification of a nucleic acid located in the region upstream (at the 5' end) of said translation initiation site.

[0069] The present invention extends to the identification of regulatory proteins that are involved in the regulation of the activity of the promoters according to the present invention. Such identification may be achieved using a yeast one-hybrid system. In such a yeast one-hybrid system the sequences according to any one of SEQ ID NO 1 to 22 are operably linked to the GAL transcription activator and transformed to a yeast cell culture. That yeast cell culture is again transformed with a library of constructs encoding candidate regulatory factors.

[0070] The present invention will now be described with reference to the following figures in which:

Figure 1 shows a general schematic representation of a promoter. Regulatory elements are sequences that may for example be responsible for special and/or temporal regulation of the promoter activity. The minimal promoter is the minimal sequence necessary and sufficient to drive expression. It includes a TATA box, which is necessary to correctly direct the RNA polymerase II to the transcription initiation site. The transcription initiation element (INR) includes the transcription initiation start site. The 5' untranslated region (5'UTR) is the region that is transcribed into pre-messenger RNA and eventually into mRNA, but is not translated into protein. The translation initiation codon is represented by the startcodon ATG.

Figure 2 is a map of the vector p4581 useful for expression in plants of a β -glucuronidase (GUS) gene under control of any one of the promoters according to the invention. This binary vector comprises a Gateway recombination cassette, suitable for the recombination cloning of any of the promoters of the present invention in front of the *Escherichia coli* β -glucuronidase (GUS) gene. This cassette contains a chloramphenicol resistance gene (CamR) and the ccdB suicide gene for counter selection of non-recombined plasmids. This GUS expression cassette further comprises the double terminator sequence T-zein and T-rbcS-deltaGA. This expression cassette is located within the left border (LB repeat, LB Ti C58) and the right border (RB repeat, RB Ti C58) of the nopaline Ti plasmid. Cloned within these borders are also selectable marker and a screenable marker genes each under control of a constitutive promoter and a terminator sequence. This vector also contains an origin of replication (pBR322) for bacterial replication and a bacterial selectable marker (Spe/SmeR) for bacterial selection.

The following figures show the results of the GUS staining of plants or plant parts transformed with the reporter vector p4581 carrying a promoter according to the present invention operably linked to the reporter gene GUS. Plants denoted "C plants" are transgenic plants grown to about 5 cm; Plants denoted "B plants" are grown to about 10 cm; and plants denoted "A plants" are grown to maturity. These A plants were used to collect different tissue samples from old leaves, young leaves and seeds.

Figure 3 shows the expression pattern of PRO0110 (RCc3, SEQ ID NO 1). GUS staining is visible in roots.

Figure 4 shows the expression pattern of PRO0005 (putative beta-amylase, SEQ ID NO 2). GUS staining is visible in seeds, more specifically in the embryo or in the scutellum of the embryo.

Figure 5 shows the expression pattern of PRO0009 (putative cellulose synthetase, SEQ ID NO 3). GUS staining is visible in roots.

Figure 6 shows the expression pattern of PRO005 (proteinase inhibitor Rgpi9, SEQ ID NO 4). GUS staining is visible in the seeds.

Figure 7 shows the expression pattern of PRO0061 (beta expansine EXPB9, SEQ ID NO 5). GUS staining is visible in young flowers of A plants (A) and in other young expanding tissues of B plants (B) and C plants (C).

Figure 8 shows the expression pattern of PRO0063 (putative structural protein, SEQ ID NO 6). GUS staining is visible in young tissues, for example in the calli (A) or old leaves, young leaves and seeds of "A plants" (B).

Figure 9 shows the expression pattern of PRO0081 (putative caffeoyl-CoA 3-O-methyltransferase, SEQ ID NO 7). GUS staining is visible in young tissues, particularly of the shoot.

Figure 10 shows the expression pattern of PRO0091 (prolamine RP5, SEQ ID NO 8). GUS staining is visible in seeds (A), particularly in the endosperm, and in meristem (B).

Figure 11 shows the expression pattern of PRO0095 (putative amino peptidase, SEQ ID NO 9). GUS staining is visible in seeds, more particularly in the embryo.

Figure 12 shows the expression pattern of PRO0111 (uclacyanin 3-like protein, SEQ ID NO 10). GUS staining is visible in roots and in meristem.

Figure 13 shows the expression pattern of PRO0116 (26S proteasome regulatory particle non-ATPase subunit 11, SEQ ID NO 11). GUS staining is weakly visible in the whole plant (weak constitutive) and is particularly visible in meristem.

Figure 14 shows the expression pattern of PRO0117 (putative 40S ribosomal protein, SEQ ID NO 12). GUS staining is visible in the seeds, more particularly in the endosperm.

Figure 15 shows the expression pattern of PRO0122 (chlorophyll a/b-binding protein precursor (Cab27), SEQ ID NO 13). GUS staining is visible in the shoot.

Figure 16 shows the expression pattern of PRO012 (putative protochlorophyllide reductase, SEQ ID NO 14). GUS staining is visible in the shoot (above-ground tissues).

Figure 17 shows the expression pattern of PRO0133 (chitinase Cht-3, SEQ ID NO 15). GUS staining is visible in the roots and meristem.

Figure 18 shows the expression pattern of PRO0151 (WSI18, SEQ ID NO 16). GUS staining is visible in the calli and upper plant parts (A) as well as in the aleurone layer and embryo (B).

Figure 19 shows the expression pattern of PRO0169 (aquaporine, SEQ ID NO 17). GUS staining is visible in the whole plant (constitutive expression).

Figure 20 shows the expression pattern of PRO0170 (High mobility group protein, SEQ ID NO 18). GUS staining is strongly visible in the whole plant as is illustrated by the "B plants" (A), and various tissues such as old leaves, young leaves and seeds (B) and calli (C) (constitutive expression).

Figure 21 shows the expression pattern of PRO0171 (reversibly glycosylated protein RGP1, SEQ ID NO 19). GUS staining is visible in all plant parts (constitutive expression).

Figure 22 shows the expression pattern of PRO0173 (cytosolic MDH, SEQ ID NO 20). GUS staining is visible in all plant parts and particularly in the shoot (above-ground tissues) and seeds.

Figure 23 shows the expression pattern of PRO0175 (RAB21, SEQ ID NO 21). GUS staining is weakly visible in calli (A), meristems and young leaves, and is strongly visible in developing and maturing seeds (B) more particularly in the embryo.

Figure 24 shows the expression pattern of PRO0177 (Cdc2-1, SEQ ID NO 22). GUS staining is weakly visible in meristem and in leaf sheets.

Examples

[0071] The promoters according to the present invention were isolated as DNA regions spanning about 1.2 kb of the sequence upstream of the translation initiation codon (i.e. first ATG, which codon was excluded) from various rice genes. For determination of their nucleic acid sequence and their expression pattern, the following procedure was followed: First *in silico* studies on genomic rice sequences were performed. However, procedures based on automated prediction programs to locate promoter-like nucleic acid sequence are highly error prone, even for the localization the best-characterized promoter control elements such as the TATA box and the transcription initiation element (INR). Also, *in silico* determination of expression pattern is extremely speculative. Therefore, to obtain unambiguous data about the nucleic acid sequence and the expression pattern of the promoters, *in vivo* studies were performed encompassing (i) Isolation

of the promoter nucleic acid sequence; (ii) operably linking a reporter gene to the promoter and introducing the resulting genetic construct into a host organisms; (iii) growing the transformed host cell under conditions allowing expression of the reporter gene, and (iv) determination of the reporter gene activity in the different tissues of the host organism. These methods are now described in more detail.

Example 1. Identification and Isolation of the promoters

Identification of rice ESTs, the corresponding genes and their location in the rice genome

[0072] Sequence databases, comprising rice sequences, were searched for rice expressed sequence tags (ESTs). Subsequently an *"in silico"* Northern-blot was performed to allow identification of EST families that are strongly expressed or that are specific for a particular organ. This analysis included normalization of the numbers of ESTs isolated from different plant organs. The ESTs families with an interesting distribution among source cDNA libraries were selected for further analysis and sequence homology searches. After sequence homology searches in combination with scanning scientific data, the genes that correspond to those families of ESTs were identified from sequence databases and a (putative) function and corresponding gene name was given (see Table 1). Subsequently, the corresponding promoter region was isolated by the following procedure. In a first step the TIGR database was searched to find a tentative contig corresponding to an EST family. Sequence homology was found using standard computer programs, such as Blast N using standard parameters (typically G Cost to open a gap = 5, E Cost to extend a gap = 2, q Penalty for a mismatch in the blast portion of run = -3, r Reward for a match in the blast portion of run = 1, e Expectation value = 10.0, W Word size = 11, v Number of one-line descriptions = 100, b Number of alignments to show = 100, Matrix = BLOSUM62). The TIGR database (The Institute for Genomic Research), provides Tentative Contigs (TC) which are sequence predictions based on contig building from all known EST, from all known cDNA and from reconstructed mRNA. The TCs used for identification of the promoters of the present invention are represented in Table 1. In a second step these TCs were used to locate the corresponding gene on a genomic sequence, which gene comprises the coding region as well as the promoter region. Generally, these genomic sequences were BAC clones, which are represented herein by their Genbank accession number (see Table 1). From these BAC clones the sequence identity of the promoter region could be determined.

[0073] Table 1: list of rice promoters of the present invention. The promoter sequences are represented herein by their SEQ ID NO and promoter number (PRO). The coding sequences (CDS) naturally driven by a promoter of the present invention are represented by their name, by SEQ ID NO and by Tentative contig (TC) accession number of the TIGR database. The Genomic sequences (BAC clones or genes) comprising a promoter region of the present invention are represented by their Genbank accession number.

Prom SEQ ID NO	Prom number	CDS name	CDS SEQ ID NO	CDS TC	BAC clone (*or gene)
1	PRO0110	RCc3	23	TC89946	AC037426
2	PRO0005	putative beta-amylase	24	TC90358	AC022457
3	PRO0009	putative cellulose synthase	25	TC83635	AC022457
4	PRO0058	proteinase inhibitor Rgp19	26	TC83117	AF044059
5	PRO0061	beta expansine EXPB9	27	TC89913	AC020666
6	PRO0063	structural protein	28	TC89985	AP001278
7	PRO0081	putative caffeoyl-CoA 3-O-methyltransferase	29	TC89891	AP000364
8	PRO0091	prolamine RP5	30	TC89670	AF156714*
9	PRO0095	putative methionine aminopeptidase	31	TC89883	AC027133

(continued)

Prom SEQ ID NO	Prom number	CDS name	CDS SEQ ID NO	CDS TC	BAC clone (*or gene)
10	PRO0111	uclacyanin 3-like protein	32	TC90434	AJ307662
11	PRO0116	26S proteasome regulatory particle non-ATPase subunit 11	33	TC83072	AP000969
12	PRO0117	putative 40S ribosomal protein	34	TC90038	AC090871
13	PRO0122	chlorophyll a/b-binding protein precursor (Cab27)	35	TC82936	AP004700
14	PRO0123	putative protochlorophyllide reductase	36	TC89839	AL606456
15	PRO0133	chitinase Cht-3	37	TC85888	D16223*
16	PRO0151	WS118	38	TC84300	AP003023
17	PRO0169	aquaporine	39	TC89687	AP005108
18	PRO0170	High mobility group protein	40	TC89846	AP004004
19	PRO0171	reversibly glycosylated protein RGP1	41	TC82935	AC090874
20	PRO0173	cytosolic MDH	42	TC82977	AC037425
21	PRO0175	RAB21	43	TC83646	Y00842*
22	PRO0177	Cdc2-1	44	TC90619	AP004765

Identification and isolation of the promoter regions of rice genes

[0074] Starting from the sequence information of the genes and their location in the rice genome, the promoter regions of these genes were isolated as the DNA region spanning about 1.2 kb upstream of the translation initiation codon (i.e. first ATG), which codon was excluded. When an intervening sequence such as an intron, was present in the 5' untranslated region of the gene, the isolated DNA region was taken as the region spanning about 1.2 kb plus the length of that intervening sequence. The promoter regions were isolated from genomic DNA of *Oryza sativa Japonica* or exceptionally from *Oryza sativa Indica* via PCR using specific primers. These specific primers comprise AttB recombination sites, suitable for recombination cloning of the isolated promoter region. These specific primers are herein represented as SEQ ID NO 45 to 88 and are listed in Table 2. Conditions for PCR were as follows: 1 cycle of 2 min at 94°C, 35 cycles of 1 min at 94°C, 1 min at 58°C and 2 min at 68°C, and 1 cycle of 5 min at 68°C. The length of the expected PCR fragment is also indicated in Table 2. The corresponding PCR fragment was purified from the PCR reaction mix via gele electrophoresis and subsequent purification using Zymoclean Gel DNA Recovery Kit (Zymo Research, Orange, California).

Table 2: Overview of the primers used to isolate the rice promoters of the present invention and the length of the rice promoter regions.

Promoter SEQ ID NO	Promoter number	Prom length	Primer forward SEQ ID NO	Primer forward	Primer reverse SEQ ID NO	Primer reverse
1	PRO0110	1264	45	prm3780	67	prm3781

(continued)

Promoter SEQ ID NO	Promoter number	Prom length	Primer forward SEQ ID NO	Primer forward	Primer reverse SEQ ID NO	Primer reverse
2	PRO0005	1215	46	prm2768	68	prm2769
3	PRO0009	1038	47	prm2420	69	prm2421
4	PRO0058	1301	48	prm2853	70	prm2854
5	PRO0061	1243	49	prm2426	71	prm2427
6	PRO0063	1019	50	prm2855	72	prm2856
7	PRO0081	1212	51	prm3025	73	prm3026
8	PRO0091	1052	52	prm3029	74	prm3030
9	PRO0095	1216	53	prm3061	75	prm3062
10	PRO0111	1237	54	prm3031	76	prm3032
11	PRO0116	1100	55	prm3051	77	prm3052
12	PRO0117	1216	56	prm3592	78	prm3049
13	PRO0122	1210	57	prm5131	79	prm2195
14	PRO0123	123	58	prm3782	80	prm2197
15	PRO0133	1808	59	prm2844	81	prm2845
16	PRO0151	1828	60	prm2973	82	prm2974
17	PRO0169	1267	61	prm3770	83	prm3771
18	PRO0170	1130	62	prm3772	84	prm3773
19	PRO0171	1230	63	prm3774	85	prm3775
20	PRO0173	1234	64	prm3776	86	prm3777
21	PRO0175	1553	65	prm3800	87	prm3801
22	PRO0177	1087	66	prm5135	88	prm5136

Example 2. Cloning of promoter-GUS reporter vectors for plant transformation

[0075] The purified PCR fragments of Example 1, corresponding to the promoter regions of the present invention, were cloned into the pDONR201 entry plasmid of the Gateway™ system (Life Technologies) using the "BP recombination reaction". The identity and base pair composition of the cloned Insert was confirmed by sequencing and additionally, the resulting plasmid was tested via restriction digests.

[0076] In order to clone each of the promoters of the present invention in front of a reporter gene, each entry clone of Example 1 was subsequently used in an "LR recombination reaction" (Gateway™) with the destination vector p4581. This destination vector was designed to operably link each promoter of the present invention to the *Escherichia coli* beta-glucuronidase (GUS) gene via the substitution of the Gateway recombination cassette in front of the GUS gene. Furthermore this destination vector is suitable for transformation of plants and comprises within the T-DNA left and right borders the resulting promoter-GUS cassette and selectable marker and screenable marker cassettes (see Figure 2). The resulting reporter vectors, comprising a promoter of the present invention operably linked to GUS, are subsequently transformed into *Agrobacterium* strain LBA4044 and subsequently into rice plants using standard transformation techniques.

Example 3. Expression patterns of the promoter-GUS reporter cassette in plants Growth and harvest of transgenic plants or plant parts at various stages (C plants, B plants and A plants)

[0077] For each promoter-GUS reporter construct, 3 T0 transgenic rice plants were generated from transformed cells. Plant growth was performed under normal conditions. The first transgenic plant was sacrificed for GUS staining when

it had reached a size of about 5 cm, which plant is named herein "C plant". The second transgenic plant was sacrificed for GUS staining when it had reached a size of about 10 cm, which plant is named herein "B plant". The third transgenic plant was kept for seed production and is named herein "A plant". GUS staining was performed on complete C and B plants. On A plants, GUS staining was performed on leaf pieces, flowers and section of seeds at various developmental stages. A plants were allowed to set seed, which seeds were used after harvest for confirmation of the expression pattern in T1 plants.

GUS staining

[0078] The sacrificed plants or plant parts were covered with 90% ice-cold acetone and incubated for 30 min at 4 °C. After 3 washes of 5 min with Tris buffer [15,76 g Trizma HCl (Sigma T3253) + 2,922 g NaCl in 1 l bidl, adjusted to pH 7,0 with NaOH], the material was covered by a Tris/ferricyanate/X-Gluc solution [9,8 ml Tris buffer + 0,2 ml ferricyanate stock (0,33 g Potassium ferricyanate (Sigma P3667) in 10 ml Tris buffer)+ 0,2 ml X-Gluc stock (26,1 mg X-Gluc (Europa Bioproducts ML 113A) in 500 µl DMSO)]. Vacuum infiltration was applied for 15 to 30 minutes. The plants or plant parts were incubated for up to 16 hours at 37 °C until development of blue colour was visible. The samples were washed 3 times for 5 minutes with Tris buffer. Chlorophyll was extracted in ethanol series of 50%, 70% and 90% (each for 30 minutes).

Expression patterns of the promoters of the present invention

[0079] The expression patterns of the rice promoters of the present invention are summarized in Table 3.

Table 3: expression patterns of the rice promoters of the present invention

PRO SEQ ID NO	Promoter number	Promoter name	Expression pattern
1	PRO0110	RCc3	strong root
2	PRO0005	putative beta-amylase	Embryo (scutellum)
3	PRO0009	putative cellulose synthase	weak in roots
4	PRO0058	proteinase inhibitor Rgpl9	seed
5	PRO0061	beta expansine EXPB9	weak in young tissues
6	PRO0063	structural protein	young tissues+calli+embryo
7	PRO0081	putative caffeoyl-CoA 3-O-methyltransferase	shoot
8	PRO0091	prolamine RP5	meristem + strong in endosperm
9	PRO0095	putative methionine aminopeptidase	embryo
10	PRO0111	uclacyanin 3-like protein	weak meristem
11	PRO0116	26S proteasome reg. particle non-ATPase s.u. 11	weak meristem
12	PRO0117	putative 40S ribosomal protein	weak in endosperm
13	PRO0122	chlorophyll a/b-binding protein precursor (Cab27)	weak in shoot
14	PRO0123	putative protochlorophyllide reductase	strong shoot specific
15	PRO0133	chitinase Cht-3	weak meristem specific
16	PRO0151	WS118	Calli + shoot + strong embryo
17	PRO0169	aquaporine	medium constitutive
18	PRO0170	High mobility group protein	strong constitutive
19	PRO0171	reversibly glycosylated protein RGP1	weak constitutive
20	PRO0173	cytosolic MDH	Shoot and seed

(continued)

PRO SEQ ID NO	Promoter number	Promoter name	Expression pattern
21	PRO0175	RAB21	embryo
22	PRO0177	Cdc2-1	weak in meristem + strong seed

[0080] The following paragraphs describe the observed expression patterns of the promoters of the present invention in more detail. The observations are based on the visual inspection of the GUS stained tissues as described above. It is to be understood that for some promoters expression may be weak and that expression in certain tissues may only be visible with very sensitive detection methods.

PRO0110 - SEQ ID NO 1-RCc3

[0081] 1 construct (OS1432), which is a reporter vector as described in Example 2 comprising PRO0110 was investigated. 25 calli, 14 C, 21 B plants and 21 A plants were analysed. There was no expression visible in calli, but strong expression in roots of C plants (93%) and of B plants (81 %) was observed. No expression in the shoots of A plants was observed. Therefore the RCc3 promoter PRO0110 is suitable for strong expression in roots.

PRO0005 - SEQ ID NO 2 - putative beta-amylase

[0082] 1 construct (OS1365) was investigated. 28 calli, 24 B plants and 22 A plants were analysed. Occasional expression in calli (7%) was observed as well as occasional weak expression in roots (4%) and shoots (12%) of B plants, expression in the scutellum of embryos of A plants (43%) and occasional expression in leaves (5%) of A plants. This promoter is therefore suitable for expression in embryo, more preferably in the scutellum of the embryo. This region of the embryo is also referred to as the transfer layer of the embryo. This promoter may have some leakiness in other tissues.

PRO0009 - SEQ ID NO 3- putative cellulose synthase

[0083] 1 construct (OS1461) was investigated. 20 calli, 20 C, 20 B plants and 20 A plants were analysed. Occasional expression in calli (20 %) was observed as well as weak expression in roots (55%) of C plants, occasional expression in young leaves (10%) of C plants and weak expression in the roots (25%) of B plants. No expression in leaves of A or B plants was observed. Therefore this promoter is suitable for expression in roots. This promoter may show some leakiness in the leaves.

PRO0058 - SEQ ID NO 4- proteinase inhibitor RgpI9

[0084] 1 construct (OS1370) was investigated. 13 B plants and 12 A plants were analysed. No expression was observed in B plants. In A plants, no expression was observed in the leaves, but there was strong expression in endosperm and embryo (58-42%). Therefore, this promoter PRO0058 is suitable for expression in seeds.

PRO0061 - SEQ ID NO 5- beta expansine EXPB9

[0085] 2 constructs (OS1441 and OS1460) were investigated. 20 calli, 32 C, 32 B plants and 32 A plants were analysed. Weak expression was observed in the leaves of C and B plants. In A plants expression in the flowers was observed (44%), more particularly in lemma of young spikelets. It was concluded that the promoter PRO0061 is suitable for expression in young tissue, more preferably in young, developing or expanding tissue, more preferably in green tissue.

PRO0063 - SEQ ID NO 6- putative structural protein

[0086] 1 construct (OS1446) was investigated. 13 calli, 13 C, 13 B plants and 12 A plants were analysed. In calli, weak expression was detected (92%). In C plants, there was no expression in roots and there was weak expression in some leaves (46%). In B plants, there was no expression in roots and weak expression in young tillers (78%) or young leaves (54%), but no expression in old leaves. In A plants, there was occasional expression in young leaves (17%) and expression in embryo and scutellum (42%). Therefore it was concluded that this promoter is active in the above-ground tissues, such as leaf, stem and seed. These data demonstrate that the promoter is suitable for expression in calli and in the shoot, and for expression in young tissues and seeds.

PRO0081 - SEQ ID NO 7- putative caffeoyl-CoA 3-O-methyltransferase

[0087] 1 construct (OS1419) was investigated. 20 calli, 20 C, 20 B plants and 20 A plants were analysed. No expression was observed in Calli. Expression was observed in C plants, more particularly weak expression in root cylinder (40%) and weak expression in young leaves (80%) and in old leaves. Expression was also observed in B plants, more particularly weak expression in roots (25%) and weak expression in young leaves (80%). Expression was also observed in young leaves (50%) of A plants. It was concluded that promoter PRO0081 is suitable for expression in above-ground tissues, preferably in the shoot. This promoter may have some leakage of expression in roots.

PRO0091 - SEQ ID NO 8- prolamine RP5

[0088] 1 construct (OS1558) was investigated. 12 C, 12 B plants and 12 A plants were analysed. Weak expression was observed in the discrimination centre (50%) of C plants and in the discrimination centre (58%) of B plants. Strong expression was observed in endosperm (55%) of A plants. This promoter was found to be useful for strong expression in the endosperm, with leakiness in meristem, preferably the shoot meristem or discrimination centre.

PRO0095 - SEQ ID NO 9- putative methionine aminopeptidase

[0089] 1 construct (OS1423) was investigated. 16 calli, 14 C, 14 B plants and 16 A plants were analysed. Some expression was observed in root-tips (36 %) of C plants and in the embryo (38%) of A plants, but not in endosperm of A plants. It was concluded that PRO0095 is suitable for expression in embryo.

PRO0111 - SEQ ID NO 10- uclacyanin 3-like protein

[0090] 1 construct (OS1421) was investigated. 22 calli, 21 C, 22 B plants and 21 A plants were analysed. Weak expression was observed in the discrimination centre and meristems (77%) of B plants. It was concluded that promoter PRO0111 is suitable for weak expression in the meristem, preferably in shoot meristem or discrimination centre.

PRO0116 - SEQ ID NO 11-26S proteasome regulatory particle non-A TPase subunit 11

[0091] 1 construct (OS1679) was investigated. 13 C, 14 B plants and A plants were analysed. Weak expression was observed in meristem/discrimination centre of C plants (38%) and of B plants (71%) and in young leaf sheaths of C plants (77%) and of B plants (21%). It was concluded that promoter PRO0116 is suitable for expression in meristem, preferably in shoot meristem or discrimination centre.

PRO0117- SEQ ID NO 12- putative 40S ribosomal protein

[0092] 1 construct (OS1425) was investigated. 9 calli, 9 C, 9 B plants and 9 A plants were analysed. Occasional weak expression was observed in roots (22%) and in young leaf blades (44%) of C plants. Expression was mainly observed in endosperm (37%) of A plants. Therefore, promoter PRO117 was found to be suitable for expression in endosperm and may have some leakiness in young leaves.

PRO0122 - SEQ ID NO 13- chlorophyll *a/b*-binding protein presursor (Cab27)

[0093] 1 construct (OS1675) was investigated. 38 calli, 38 C, 38 B plants and 15 A plants were analysed. Very weak expression was observed in the discrimination centre and young leaf sheaths of C plants. It was concluded that this promoter PRO0122 is suitable for weak expression in shoots.

PRO0123 - SEQ ID NO 14- putative protochlorophyllide reductase

[0094] 1 construct (OS1433) was investigated. 21 calli, 18 C, 19 B plants and 18 A plants were analysed. Strong expression was observed in shoots (33-68%) of C plants and B plants (63-79%). In B plants there was also occasional expression in roots. In A plants, again strong expression in young leaves (73%) was observed, as well as occasional expression in old leaves (39%). It was concluded that this promoter is suitable for strong expression in shoots, preferably in leaves.

PRO0133 - SEQ ID NO 15- chitinase Cht-3

[0095] 1 construct (OS1687) was investigated. 15 calli, 12 C, 16 B plants and 12 A plants were analysed. Weak expression was observed in calli (66%) and in the discrimination centre/meristem (50%) of B plants. It was concluded that promoter PRO0133 is suitable for weak expression in meristem, preferably in shoot meristem or discrimination centre.

PRO0151 - SEQ ID NO 16- WSI18

[0096] 1 construct (OS1458) was investigated. 22 calli, 16 C, 16 B plants and 13 A plants were analysed. Strong expression was observed in calli (91 %) and weak expression in shoots of C plants (62%). In A plants there was very strong expression in the aleurone layer and in the embryo (46%). It was concluded that promoter PRO0151 is suitable for strong expression in calli and in seeds, more particularly in the aleurone layer and in the embryo of the seeds.

PRO0169 - SEQ ID NO 17- aquaporine

[0097] 1 construct (OS1911) was investigated. 11 calli, 10 C plants, B plants and A plants were analysed. Some expression (55%) was observed in calli and in roots (30%) of C plants. Furthermore, good expression was observed in shoot tissues (80%) of C plants and in young leaves of B plants. It was concluded that this promoter is suitable for constitutive expression, preferably constitutive in young plants.

PRO170 - SEQ ID NO 18- High mobility group protein

[0098] 1 construct (OS1434) was investigated. 23 calli, 21 C, 21 B plants and 14 A plants were analysed. Expression was observed in calli (52%) and in roots (51%) of C plants. Moreover, strong expression was observed in young leaves (81%) of C plants, in roots (86%) of B plants and in young leaves (86%) of B plants. In A plants there was strong expression in young leaves (75%), old leaves (43%), embryo and aleurone but a weaker expression in endosperm (82%). It was concluded that promoter PRO170 is suitable for strong constitutive expression.

PRO0171 - SEQ ID NO 19- reversibly glycosylated protein RGP1

[0099] 1 construct (OS1762) was investigated. 18 calli, 11 C and 13 B plants were analysed. Strong expression was observed in calli (44%) and in all tissues (27%) of C plants. In all tissues of B plants (16%), expression was somewhat weaker but most pronounced the in discrimination centres (46%). It was concluded that promoter PRO0171 is suitable for constitutive expression.

PRO0173 - SEQ ID NO 20- cytosolic MDH

[0100] 1 construct (OS1435) was investigated. 17 calli, 17 C, 17 B plants and 15 A plants were analysed. Occasional expression (12%) was observed in calli and weak expression was observed in upper parts (24-69%) of C plants as well as in young leaves (41 %) of B plants. In A plants, expression in leaves (33%) was observed and strong expression in seeds (38%), but not in the root. It was concluded that the promoter PRO0173 is suitable for expression in above-ground tissues especially for constitutive expression in the shoot and especially in the seeds.

PRO0175 - SEQ ID NO 21- RAB21

[0101] 1 construct (OS1436) was investigated. 16 calli, 12 C, 15 B plants and 15 A plants were analysed. Expression was observed in some calli (31 %), in the discrimination centres (42%) of C plants and in young leaves (25-58%) of C plants and A plants (15 %). Furthermore, very strong expression was observed in aleurone and embryo (60%) of a plant. It was concluded that promoter PRO0175 is suitable for strong expression in calli and in seeds, more particularly in developing/maturing seeds, more particularly in the aleurone layer and in the embryo of the seeds.

PRO0177 - SEQ ID NO 22- Cdc2-1

[0102] 1 construct (OS1436) was investigated. 16 calli, 12 C, 15 B plants and 15 A plants were analysed. Expression was observed in some of the calli (31%), in the discrimination centre (42%) of C plants, in young leaves (25-58%) of C plants and occasionally in young leaves (15 %) of A plants. Moreover, very strong expression was observed in aleurone and embryo (60%) of seeds from A plants. It was concluded that this promoter is suitable for specific expression in seeds, more particularly in developing/maturing seeds.

Example 4. Stability of the expression patterns of the promoters of the present invention in further generations

[0103] The above-mentioned analyses were performed on T0 plants originating from the transformed tissues. The stability of promoter activity in the next generations or progeny plants of the original T0 plant, the so-called T1 and T2 plants, was evaluated as follows. The T0 plant transformed with the reporter constructs as mentioned in the above paragraphs of Example 2, were grown until maturity (A plants), of which the seeds (T1 seeds) were harvested and sown to generate progeny T1 plants. These plants were analysed as described above in Example 3 and the A T1 plants were allowed to reach maturity and to set T2 seeds.

[0104] The expression pattern of the promoters of the present invention was studied in T0 plants, T1 seeds, T1 plants and T2 seeds and in all the tissues (including seeds and seed tissues) as described in Example 3. The specific expression patterns as reported from the T0 and T1 seeds and described in Example 3 were confirmed in the following T1 generation and T2 seeds. It is concluded that the expression pattern of the promoters of the present are stably inherited in plants of subsequent generations.

Example 5. Stability of expression patterns of the promoters of the present invention in other plants

[0105] The above-mentioned plant analyses were performed on rice plants. This choice was based on the practical consideration that plant genetic engineering is most profitable for crop plants. Also in other crop plants, such as for example *Zea Mays*, the reporter constructs comprising the promoters according to the present invention are introduced and transformed plant are evaluated as described hereinabove. The expression patterns of the promoters according to the present invention are conserved among plants. Therefore, the promoters according to the present invention are also suitable for driving and/or regulating expression of an operably linked nucleic acid in monocots, such as corn.

[0106] For many other purposes such as research and horticulture, (small) herbs are being genetically modified, which involves the use of promoters. Therefore the reporter constructs comprising the promoters according to the present invention are introduced into other plants species such as for example *Arabidopsis thaliana* and transformed plants are evaluated as described hereinabove. The expression patterns of the promoters according to the present invention are conserved among plants. Therefore, the promoters according to the present invention are also suitable for driving and/or regulating expression of an operably linked nucleic acid in other plant species such as for example dicots, such as *Arabidopsis*.

EP 1 801 223 A2

SEQUENCE LISTING

5 <110> CropDesign N.V.

<120> Rice promoters

<130> PF58322

10 <150> EP 03075331.3

<151> 2003-02-04

<160> 88

15 <170> PatentIn version 3.1

<210> 1

<211> 1264

<212> DNA

<213> Oryza sativa

20 <220>

<221> misc_feature

<223> PRO0110 - RCc3

<400> 1

25 tcgacgctac tcaagtgggtg ggaggccacc gcatgttcca acgaagcgcc aaagaaagcc 60

ttgcagactc taatgctatt agtcgcctag gatatttggg atgaaaggaa ccgcagagtt 120

tttcagcacc aagagcttcc ggtggctagt ctgatagcca aaattaagga ggatgccaaa 180

acatgggtct tggcggggcg gaaacacctt gatagggtgg ttacctttta acatgttcgg 240

gccaaaggcc ttgagacggg aaagttttct atttgcgctt gcgcattgtac aattttattc 300

ctctattcaa tgaatttggg ggctcactgg ttcatataaa aaaaaagaat ctagcctgtt 360

30 cgggaagaag aggatatttg tctgagagag gagagagaga gagagagaga gagagagaga 420

gaaggaggag gaggattttc aggtctcgca ttgcccaccc tctgcttctg ttggcccacg 480

aagaatccca ggcgcccatt ggctggcagt ttaccacgga cctacctagc ctaccttagc 540

latctaagcg ggcgcacctc glagccacgt gcctagtgtg gattaaagtt gccggggcag 600

cagggaagcca cgtgcgaatg gcattctccc ctgtctctcg cgtacgtgaa aacaaaacca 660

35 ggtaagctta gaattctctt gcccggttgg ctgggacacc caccaatccc accatgcccc 720

gatattcttc cggctctcgg tcatgtgatg tctctctctg tgtgatcag gagcaagcat 780

tcttaaagcg caaaagaaaa tcaccaactt gctcagcgag tcacgctgca ccgcgcgaag 840

cgacgcccga taggccaaga tcgcgagata aaataacaac caatgatcat aaggaaacaa 900

gccgcgcatg tgtcgtgtgc agcaatcttg gtcatttgcg ggatcgagtg cttcacagct 960

40 aaccaaatat tcggccgatg atttaacaca ttatcagcgt agatgtacgt acgatttgtt 1020

aattaatcta cgagccttgc tagggcaggt gttctgccag ccaatccaga tcgccctcgt 1080

atgcacgctc acatgatggc agggcagggt tcacatgagc tctaacggtc gattaattaa 1140

tcccggggct cgactataaa tacctcccta atcccatgat caaaaccatc tcaagcagcc 1200

taatcatctc cagctgatca agagctctta attagctagc tagtgattag ctgcgcttgt 1260

gatac 1264

45 <210> 2

<211> 1215

<212> DNA

<213> Oryza sativa

50 <220>

<221> misc_feature

<223> PRO0005 - putative beta-amylase

<400> 2

55 cccgatttag tagaccacat ttggcatca aacaaaata gacctctccc cagaatttgt 60

EP 1 801 223 A2

	aatgggcttt	gtgggttcgtg	atatcactga	acctgctggg	tgaataaagt	aaaaaaaaa	120
	accataaat	tggccttctg	caagatctcg	tcgtcttgcc	caaaactatag	ccttcgatct	180
	ttccatcagg	accgcatggg	gggagagcag	gggcaagtat	gaaatggagt	tcagattcag	240
5	attctagaac	agtctgaaca	tgcgacgacg	acgatggcga	tgtatctgaa	caatctggtc	300
	ctctccctct	cctcccgggc	gggcttccac	gcggctgagt	ttcaggctcc	caatctgcag	360
	ctcctcccag	aaccttactc	tgattgattg	gttcacogtt	tccatggctc	caatgaatgc	420
	aacgtgttgt	tcagattttc	tgaatcttgt	tctcaatccg	gagtacgtgc	tgtagcagca	480
	gcaatctgtc	cctgatctga	gaattttaga	cactcgtaga	ttcgctgatc	aatcattccg	540
	tcccttcgag	tggtctagat	tgagcttaat	catcctgcta	ctcgaatcaa	atcttcagca	600
10	agtggagagct	agataattca	gaagaaatca	acataattctt	cgcgaaaaaa	agaaataacc	660
	gatgaaacca	cggtaattag	gttcttcgaa	tcaccgggag	agtaggaaaa	aacgagctaa	720
	aatcccacat	aggaggaaac	ggttaaaaac	ggccactccg	cgtctccgcc	gcgagactag	780
	ctctcgccag	tcacgtagc	ccaatccaca	accgccacgt	gctccgacaa	tcccgcccgt	840
	ccatcgccgc	ggccccggcc	tcatctcgac	cactcgtttc	ctcccttcac	accagccacg	900
15	tggcactctc	tcgagagctc	ccgcccgcct	atataaaactt	gttcgcgctc	ggctcctcct	960
	cctcctcgag	ctccacccca	cattgaataa	ttatttttaa	taatttttagt	tttttttttg	1020
	gcttttagata	tattcccaat	ccccaacctc	ccaataatcc	gatctctccc	agttctgttc	1080
	ggatcaaggc	tgtgtcgatc	gcaaaaaaga	aaaaaaaaac	aatttccttt	tgggggtggt	1140
	catctgttga	tcacttcttt	gtttcccgcg	ttttgttggy	gattcgattt	tcgggttaag	1200
20	atthttctaca	cgacc					1215
	<210>	3					
	<211>	1038					
	<212>	DNA					
25	<213>	Oryza sativa					
	<220>						
	<221>	misc feature					
	<223>	PRO0009 - putative cellulose synthase					
30	<400>	3					
	gccatcgagt	ggtgtgccga	taccggcgcc	tgttctttac	agcctcagct	agtgtttgtg	60
	tccgaggcaa	tttttccgac	ctattgtgtt	gctttcctct	ctgatagctt	atggtaaaag	120
	atacaaagat	gttgaggagt	ttgtacgcca	cttaattttg	ctcgtaacat	acattgacaa	180
	tcaagaggag	ccatggcatt	gcgatctgct	tacacggcat	attcttactg	gatgggtgtac	240
35	actacttacc	ctttttaatg	caagcatcaa	tccattgctt	ttctcactgc	acacctgatt	300
	cgtactgaaa	acgtgaaaca	taaaaaaaa	acaaaaatct	agctgatgtt	ggctctcggg	360
	gctcagagtc	tagtttgttc	tagatggcta	acctgatatg	tgttggtcac	gctcacgttt	420
	gaaccgagaa	agagtgtgtg	tgtgtgtgtg	tggcggtgct	gctacaccag	agcctccctg	480
	aatcgcaatg	cgtgttaacg	ccagcatcgc	aggatttcat	ctcacttgac	aggttcagat	540
	ggccttctct	ctaccgtctg	ccattttatac	acgcagtgc	ttaacgctta	cacgagccgg	600
40	atggcccggg	tctccccctt	gcaccatctc	accagaaaaa	cggtagaggc	tcaccgcaac	660
	ccacccacca	aacacatcca	cgtcccttca	ccgttggcct	tcgattttgc	ttcagctgca	720
	ctacgacccc	tccaacacat	ttccctcgcg	tctcgttgog	atctcacctt	acgacgatct	780
	cgttcacgca	gcagcagcat	cggcagcgcc	ggcttgcttc	cgaagcgagc	aatgcatggc	840
	gcgcgcggcc	gcgtgcgtgc	gtgccttggc	ttgcgctcta	atcaaacagg	gacgccccaa	900
45	ctcacgggtg	gtgcgggacg	ccaccccggc	accttaacgc	ccccgcctcc	ctgcatctga	960
	tcatcaacca	gctgctatat	cacctagcta	gccgcgcgct	cctcctcgcc	caccaacgtc	1020
	gcttccccgg	cacctcac					1038
	<210>	4					
50	<211>	1301					
	<212>	DNA					
	<213>	Oryza sativa					
	<220>						
	<221>	misc feature					
55	<223>	PRO0058 - proteinase inhibitor Rgp19					

EP 1 801 223 A2

	<400>	4	
	tctcttctga	agctgaagcc	ctgcgaaata
	atcatatcgg	cgtaagaccg	gtttaaacat
	cgacgtgaca	catggcacgt	ccttggagct
5	gccgagtggg	tgagcggcca	aatccgtttt
	ttgatactaa	taccatttta	tcaagcagta
	actgtgagat	ttcatcccgt	ttcaagatga
	ggcagaagac	agtggazaga	caggacaaat
	acggtttctg	aaagttgaca	tggaaatcgc
10	tctgcgattt	ccactttcgt	ttgctttcgg
	catcgtgctt	tggatctcag	caccgtagta
	aataacatgt	tctttttatg	ctgctttaca
	tcataatatta	tgagatggag	ggagtaaact
	catcaactaa	ttaaagcaca	agatacgcac
15	taaagaagat	gtcgcgcgta	acttagataa
	ttattctctc	cggaaattaa	aagctagttt
	atagtaagcg	aagcatacct	gacgtgttag
	gctagatata	gacagacact	cctcgtgatg
	aagcagctgg	ggatcgatcc	atccatccat
	cactgtcgac	tcttgcatgt	ttgcatgtaa
20	cagtgccttg	cagctgcttt	ccaaaatgct
	aaaaaaagta	gtgtgttctt	gagcctatat
	ccatcgatcc	atctgacggg	tagttccaag
			ggaaagaaga
			a
	<210>	5	
25	<211>	1243	
	<212>	DNA	
	<213>	Oryza sativa	
	<220>		
30	<221>	misc_feature	
	<223>	PRO0061 - beta-expansin EXPB9	
	<400>	5	
	aaaaccaccg	agggacctga	tctgcaccgg
	gttttccgat	cgagggacga	aaatcggatt
35	ttattccgga	gcatgattgg	gaaggagga
	gtccagatct	ccagatcact	cagcaggatc
	ttcggcttcc	cgcaaggcgg	cggccgggtg
	gcacgcggcc	gccgcggacc	cggctctgcg
	atagatagct	actactctct	ccgtttcaca
40	atattgatgt	taatgaatat	agacatatat
	atgtaggaaa	tgctagaatg	acttacattg
	tggtatggatg	caggatcatg	aaagaattaa
	ttactaattg	cgctgcata	atgcatgaca
	ccattaggaa	gtaaccttgt	cattacttat
	tcattgagcaa	atctacaaaa	ctggaaagca
45	cattaatcac	caaatatttc	gccttctcca
	tgtacacact	gacagtgtac	gcataaacgc
	gcacactggc	cttccatctc	aggctagctt
	gcgcgcgcac	aggcacaaat	tacgtacaaa
	gaatcgctcc	cgcgcgcggc	ggcgcgcggc
50	ccacgacacg	atcgcgcgcg	acgcggcgga
	ccgactataa	atacgtaggc	atctgcttga
	aggaaaaaaa	aacaaaacac	accaagccaa
			ataaaagcga
			caa
	<210>	6	
55	<211>	1019	
	<212>	DNA	

EP 1 801 223 A2

<213> *Oryza sativa*

<220>

5 <221> misc_feature

<223> PRO0063 - structural protein

<400> 6

cctagctata	tgacagaggtt	gacaggttgt	ctcttagatc	gattaataat	atcacattga	60
tgcaattaat	tatctgagat	caataaagtt	ttctcttatg	ttaaattaat	atcagtaata	120
gatgctaagt	ccttcattag	tagtatccca	catttaatac	cagttggaca	cacaaaaaaa	180
aaggcaatgc	cattaatatg	ccatctctct	tgttttccat	tgccatccaa	gtgccatatg	240
atatcatcat	caggcacacc	aatccataac	tagttcatta	gagcaagttt	aataatagag	300
ctaactataa	gcttataatt	tatatggag	taaacatgta	tagtaaatga	gctataaggt	360
tatttctttt	tttctcctcc	tctctctatc	tcttacctat	atatttaatg	tatttgtctt	420
gaagtatgtg	aatagctagc	tcttgtatga	gagccaatcc	tctgcatttt	ttaaattctc	480
tttctctcac	ataagcatat	agttggctta	tagcctgcta	ttatacttgg	tcttagtaca	540
ctaaccoccc	ttacatgcaa	tgcaagctgt	ctaattaaaa	gggtttcaca	acattttgaa	600
tgccactact	agctcccaac	cacaaccaca	gatctagcta	gggtttgttc	atttctctcc	660
tctctctccc	tcctccttcc	cgttgtgcca	attcatccaa	agtcattgag	agccatacta	720
ctccatatac	tattactcct	acatgtgtac	tacatttata	ttgatgatct	gtaagagcaa	780
aagtattaat	ggggatcaca	ggattgcagt	aacagcagca	ggtacccctc	cctttaacat	840
ccgcagttac	gcctccacc	taccgtcttc	tctgcgcatc	gatgacgatg	agcttctcct	900
ccgctataaa	tcctctcccc	tcctctctcc	ctcctcctcc	aactccacat	cgatcagcag	960
cagcagcagc	ttgcacactc	gagcttagct	tagcttttgc	aagagagacc	gagctagag	1019

25 <210> 7

<211> 1212

<212> DNA

<213> *Oryza sativa*

30 <220>

<221> misc_feature

<223> PRO0081 - putative caffeoyl-CoA 3-O-methyltransferase

<400> 7

atggtgccat	gtcaataaga	catcataata	gaaactacac	tcacacaacc	atagtttctt	60
aaagtgggtc	attaataaat	acatcatcta	tcttttctat	caatcatatt	tattctttat	120
ctattatgac	ggcactatct	tctcccaatg	taaaacctga	taatgtctag	tgcatagggt	180
ctcgtgttga	agctgtttct	tacatgagac	ccagtttctt	cttctctcca	ctctctctta	240
attaatataa	tgtcacataa	gttaaaagtt	ctagtaaata	ataatatagt	taatgacata	300
gacaacatcc	tagatgtagg	gttaggagtc	ttcggacagt	agcaaccctg	ttttgactcc	360
tttttttggt	gcccattccc	agtcgccacc	agaaaattca	ctgtgcccaa	atcaatggaa	420
ggcgcctact	gatccatcca	tcttcgtgac	agctccgagc	tttctcctgg	ttatttttct	480
cccaaaaata	cattcagaac	acgatctcaa	atttaaacta	atggagtgtc	actgcatttc	540
ttaattataa	gtcgagcac	cactcattaa	tcatttccat	cacaggtaaa	tcgtgggtgag	600
ctggtggttg	ctactgtact	actagtacta	cctgtgcgag	ctttgtagaa	gccgttttcg	660
ctgaagcttc	ttcttcttcc	ctgggcaaaa	taattttaag	caggcggaat	aatattggga	720
taaacagggt	ggacaaaagc	gtgcgatccc	tttctttaac	caaaccacga	cgaagcagg	780
ttaggtcgcg	gcagggtggtg	gtggtaggaa	gaagaagaaa	gagaggggaa	aaaaaacaaa	840
aatttcacat	gcacatgca	tgaagtagta	catgtagtac	tgagtactgt	aataatgttc	900
agtctactgg	accgtctcaa	cggaagacc	aaattaacgc	ttataaaata	cccttttttt	960
gggcactgat	catggccact	acgtttggtg	gctcaacaac	caggtcacccg	tgcatcgat	1020
cgattgctaa	tttatttttt	gaaaagggaag	ggaggaaaaa	agaccgggtg	tttggtggcg	1080
ccaccaaccc	tgctctcgtg	agccgataaa	tattgctcgc	cggagctctc	ggttgacgac	1140
ccaaccaatc	gactcgacc	accaccagca	gctcaagcag	caacagctca	aacggaggaa	1200
gatctcatcg	cc					1212

55 <210> 8

<211> 1052

EP 1 801 223 A2

<212> DNA
 <213> Oryza sativa

<220>
 5 <221> misc_feature
 <223> PRO0091 - prolamine RP5

<400> 8
 10 gtttttctat gaaccgggtca ttaaaccggtc cccgggttaga ccgaacaagc cacaataatc 60
 ttgaaatggg ccttgatgtg gcccaattgg tctgcctaga gcgttttggt tggcaaaaat 120
 caatctccta ttctcggcac gtgtgatata caatggtaag tgagatatac aattctcggc 180
 acggctacat tacaaggtgt cgcattgtgt caatgtttgg ttaatttgct agattcacat 240
 aatacatgcc aggaagttca gaacaatgtg ttgcctttca ccggaact ttgttgagc 300
 aaatgccttc ttcttttttg cttctgcttc ttgagtcctat gtggaggaag cagtagatag 360
 15 ctgatgatat caggattcct tctgtgtctg tgtaggtgta gcaacaccac tataattttt 420
 atttagcaac acaatatcaa ttgtgtctat aaaagtatga attaaatcaa tcccaacca 480
 caattagagt aagttggtga gttattgtta agctctgcaa agttaattta aaagttattg 540
 cattaactta ttctgtatca caaacaagtt ttcacaagag tattaatgga acaatgaaa 600
 ccattgaaca tactataatt ttttttctta ctgaaattat ataattcaa gagcataaac 660
 ccacacagtc gtaaagttcc acgtgtagtg cattatcaaa ataataagctt acaaacata 720
 20 acaacttag ttcaaaagt tgcaatcctt atcacattga cacataaagt gagcgatgag 780
 tcatgtcatt atttttttgc tcaccatcat gtatatatga tgggcataaa agttactttg 840
 atgatgatat caaagaacat ttttaggtgc acctaacaga atatccaat aatatgactc 900
 acttagatcc taatatagca tcaagcaaaa ctaacactct aaagcaacog atagggaac 960
 atctataaat agacaagcat aatgaaaacc ctctcatcc ttcacacaat tcaacatta 1020
 25 tagttgaagc atagtagtag aatcctacaa aa 1052

<210> 9
 <211> 1216
 <212> DNA
 30 <213> Oryza sativa

<220>
 <221> misc_feature
 <223> PRO0095 - putative methionine aminopeptidase

<400> 9
 35 cctgatggat gatgaatcac tgatcgattt ctagttctta ttctctgaag atgaaccgaa 60
 gatccaagat tggatcatga aattatcctt tcttgatttg gccctccgag aatagattcc 120
 tgtgcaatct agtcagtagt tgttcaggtc atgtaaacgt acggtaagaa atttatg-cg 180
 agagggtttt ccagtttata ctagtcattt gacctctggt catgtattga ttctgagaca 240
 40 aagtgtagtg atcgcttgat gatactagta cacattgctg ccttcttttt tgtcctgtaa 300
 aagatttatt attggcagca atggatggtg gagagggcaa tctgcttctt agttttgagt 360
 ataaaagtttt aagttttgag cagagtttgc aaaatttgca gtagaaaagt tgaaatttca 420
 aattggaagt acagtttttc aaatttccag tataaatttt taaaccact gagaaaccaa 480
 gagcatatgg gcgatcaaaa atttcttttc taaaggaaaa atatttttta aaaaacactt 540
 45 agtagtatat caaaattctg aggttaagctc attaggccca ttcactgtac ggcccatgaa 600
 gccacgtctg gtgagatggg cctaccctgt caggcagaga tggatgggccc tttaattgta 660
 ggcccatggt ggaaagccca ccaaagccca ataatatatc ctctcacct tcaaccctaa 720
 tctctctctt cttctagaag actgaaaatt cctctcctt cttctctgc cctaccgct 780
 cgccaggtt gccgtctcct tgtctcctcc gctccttgcg ccgccgccgc gacgagtcgc 840
 50 ggggaggggc ggcgatctcc atctccatct gaggcgagga gagcagggga ggtgagggga 900
 tcttggtgag gtgagcatcc acgtcctctt tcttttttct tgattcatct ctctctctct 960
 cgcacatcgg gactggaatt tgcttgctt cgttcgttaa gtttaacccta gcttctcttc 1020
 tagatctgga agaaactctt cttcttttaa tttcagagcc ttaaccttaa tagtacaagt 1080
 aacagtttgt ttgttccccg aaaagtttg atgccttcca aatagagaca catgttattt 1140
 55 attttggagt gtaatttgc cctggattta ttcattcagg tttgtgatta ctggacaata 1200
 gaaatattta cacaat 1216

<210> 10
 <211> 1237
 <212> DNA
 <213> *Oryza sativa*

5

<220>
 <221> misc_feature
 <223> PRO0111 - uclacyanin 3-like protein

10

<400> 10
 tcgttaagtt tgatgatttc tgatgaccca tggtcaccta ggggctagca gtaccatgca 60
 tgatcacccct ccacaaagaa atgggtacagt acatctccgt cccaaaataa gtgcagccat 120
 gtatatcccat gcctaacggt tgaccgtccg tcttatttaa aaaaattatg aaaaatttaa 180
 aatatatttag tcacacataa agtattattc atgtttttac atctaatagc aacaaaaaat 240
 actaatcata aaattttttt taataagata aacgggttaa cgttgaacgt gaatagtgca 300
 15 aaacttattt tagaacggag ggagtacgaa gtaactccgg aactacatat agggcaatta 360
 ttgccctatg tatgcatata gtcaatcaat taactgctga caatggaaaa gctaataaat 420
 caatcaatgg tttgattaat caaattaagc caggtcagtc cgtcagtgta cattcactaa 480
 ttaaattaac aggtttgttc aacggttcaa ccaacatctg ccatcaacat cttttcgttg 540
 cacctttctt gactctttat gctattttgc taaaaaaaaa cttctcttta catcacttat 600
 20 aacaatatat atttctgctt taatttgtaa tctttttttt ctgcggttgca acggaaatca 660
 cgagcgatat atgggtgaaga ctgatgataa tegtatttct gatgacccat gattccgcgg 720
 tgtaccatct gtctctgcaa ctaaaaagtg gagtagttcc ttgacggaag aagggaagcaa 780
 aatagaagat attctcagtt gatctgcagt tgttggttagg tcaatatatt cagaaatcgc 840
 agttgctggt gtttaaattg tgtgtgacag cagacagcta attatcagta cagctatatg 900
 25 agcaatacta gtgaatctgt actaatttaa cgagagtatt ttctatatatc aaatacaaca 960
 gcaaaactgt gccactggcg ccgaatacgt acggacagag ctcaggcaat caggggagca 1020
 gcaaaagagg agagagttgg tgccaagcac aactaaaccc aactgcaccc aaaaactaat 1080
 cagcatttca gttcgcttta gttagtacta ccacctgcac ctctttacca acatatata 1140
 accgcagtg gacctgcagt catctcacta attcagtgaa gccaccagta ctagtacggc 1200
 30 tctaatacgt tcgcgtttgc taattaactc tgccatc 1237

<210> 11
 <211> 1100
 <212> DNA
 <213> *Oryza sativa*

35

<220>
 <221> misc_feature
 <223> PRO0116 - 26S proteasome regulatory particle non-ATPase subunit 11

40

<400> 11
 ctaagggcag cagccattgg gctctatagg tgtgggttga agtgcaacta caagcgagca 60
 acctggtaga atatccccga gatcagtagt taccgtgatt gggtcagact tgagaggcta 120
 attttttcgt acctgtagct ttattacatc gcatttcctc ttattgaagt ttagccgagg 180
 tggtagggat ggatattcag tctaacagac tcaatgaacg ctttggttga tgacttgtag 240
 45 agtactggct gctcgaacag gatggttcag cttccagaaa tttggcaacg ctccatttca 300
 aagaaaaatca ttcagtatct gcttctctgt tgttacattg atctcatata aagtcacttt 360
 gatcgttgac atcttgtttt ttggttcggt tgccatggtg gtttcccttg ctgctgggag 420
 gattgccgcc tgaacttttt cttttttgag aggatgttat ttttgccaga caagaacggg 480
 aataagcaaa ttgtttggtg gaactaaagt aaactcgatc tctttccgag aagtgtatta 540
 ttttcacgtg taccatcaat ttttttgaag gtaaatattt ttccccttta actaatgttc 600
 50 actttggacc ggataatctt acctttattt aactttgggc tatstaactc tcttctaaag 660
 catataaacg atcttgagta catcgattcc tacttatcat ttaactctcg tagcttaatg 720
 taagattatt tctttgaaat atgataaatt ggatgcataa gaatgaaaga gtcaaggatt 780
 aagtgtattc tcaaaaaaaa aaaaagagtga aattttattt tttttccctt ttcgacacga 840
 agaagggctt gggttgagga aaatggccca gattcagatg accgaggccg agtaccatgg 900
 55 ggcccacaag aataataagc cccgagccca aacgctaagg cccacgagaa gccgtgcgct 960
 ggaagaaaga aagaaaccgc ggccgtcttc acaccgaagc ggcggacgag acgactcgca 1020

EP 1 801 223 A2

gtcgcagcct ctttctctct ccgtctctct ctcccccttt cctctctctcc gcgcggcgaa 1080
cgaagcgagc gagcgcgagg 1100

<210> 12
5 <211> 1216
<212> DNA
<213> Oryza sativa

<220>
10 <221> misc_feature
<223> PRO0117 - putative 40S ribosomal protein

<400> 12
15 cgtgttcatg ttccgcattha ggattggact tttttaggat ggagaggata tgtcctaacg 60
gaaatgtcat gtctatgtct cgatcttata aatttgttca atagcgttgc aaacgcgac 120
attaaaaagg cggtaagaga actaccacat tttcgaaagc ccattctctt cgtgagttac 180
tggaattatt tggcatagca catgcataaa gatgcttttag taatgagctc aataaaacac 240
gacagctttg catgtagcca caatgctata gtaaagtgtg tgtacttctt ttgcattgca 300
aagtggtagt gaccttgttt aggcagctag ctccattcat tttttgaatt ctatagttat 360
20 agttataaag attatcataa tttagataag aatccgggat gtttgagaag ctggagtttc 420
tagagaagct ataacaactc gaagctccct aaacagagcc attgaacatt gagctgtcca 480
gtatatcatg acaaaatgat acatttttga tgggcatatg tgtctaagaa aacaaacatc 540
acaattcaat cgttcaactct aaaaaaaaag gcaaaaacact caacaaaacc ataccgtgaa 600
agtgaacctc taatgaaatg aaattttgat aagcatgctt acccaggttg aaatttcaat 660
ctaagaacaa ttccaaaac caccgtccat agaaatatgt ggaattcatt cagaattttc 720
25 ataccacacg ataaaattta tagggaattt aacttttgcc atttttaccg aacaccacct 780
tttcatttgc tctataatg ttatcgaaaa gagagtgttt gttaattatt tgtcactttt 840
atcacgacat gtacccgtga caacgtggcg ttctcgttgg agcccacccg tcagccgcgcg 900
tacgcaccac catcaaagaa ttcaagacgg agagcgtcgt cgccgtcggc aaggcggcgt 960
gttttttcca cgttacgttg cttcggcgtg ggcccaatct tgttcggggc taactagtct 1020
30 ttcccagccc agggccatta agcctaccaa cccggacggc ccgggaggag ctagggtttc 1080
acccttcaat atataaacct ctctctctc ctccggcgcg cgctccgaa gccctagctc 1140
ctccgcgcgc cgccgcgcgc gccgcgcgcg cctctccact cgagagaccc agccgcgcgc 1200
gccgcgcgcg ccgcca 1216

<210> 13
35 <211> 1210
<212> DNA
<213> Oryza sativa

<220>
40 <221> misc_feature
<223> PRO0122 - chlorophyll a/b-binding protein precursor (Cab27)

<400> 13
45 cagatgccac agtatggtgt accaccagct gctccacacc atgctccacc ggctggccaa 60
ccaatgtatt tcccgaata atctatcttt atccgatgta caagcaatta gagcaattgc 120
aaatgtttgc tgcaatactc gggctctgggt atcttctctt caaatttttg gttgtaactc 180
gtctatgcag ctattcatat tgtaactcag tgagctccct gtcgcaaatg tgcctctgcg 240
tcagtcgctg tctgtaaact gtccggcaat tagaaattcc catccttagc atgcctgta 300
ttgttcagct cgaaaactgaa atttttcttc gtgccctata ttttttcggg ttagataagt 360
50 gttccgctgg aattttatgc aggtgctgta ccctatgtgc tgcctttttt ttgtgtgggg 420
cgcccccccg gggggggggg ggggtttcct ggcatgattg caaataagaa ccccggggca 480
aatctgctgg ttggttgcaa ataataaccc ctccaaatct gcgcagatga aacccattc 540
aggacatgaa ttacgattgt tcatgagcta tttggatcat ggaaagattg gaaacaaaca 600
cttacgtcaa ggtttctact aattacgtga ttccgatttc agagtcagcc atggctatac 660
tgccttttgt ccagttaaaca tcgctgctct agtaacaaac attgcagtaa acatcacaac 720
55 tatccaattc ccttgttgc gctctagtaa aaaacattgc aattatccaa ttcccagata 780
ttttctttca ctactccaaa acctaaagta catatacgtg agttgagtga tccagcaaca 840

EP 1 801 223 A2

taaaaatccg aggcctccgag cgatctgcac caaccatctc acccgctccga cgtggcagca 900
 gcaaccagcc acagctgaga cctccatcca atagaaaccc tccctttgat tcccccgat 960
 cccggcatcc ggataacgct ggataagagg cgacgcctcc cattggccac acccaaccaa 1020
 caacgcaccc tggccgtccg atccaccccc accgcgcatc tcgcgcgtcc gtcgcgcgcc 1080
 5 tcgccaccgt ggccacctgg cagcgccggc cactcccgga cagtttaata caagccacgc 1140
 ctttgctccg tgcgggcca aacgtaccct tgtgactaca cccgcttcgc ttcctccct 1200
 ctctaagccg 1210

 <210> 14
 10 <211> 1179
 <212> DNA
 <213> Oryza sativa

 <220>
 15 <221> misc_feature
 <223> PRO0123 - putative protochlorophyllide reductase

 <400> 14
 ttgcagttgt gaccaagtaa gctgagcatg cccttaactt cacctagaaa aaagtatact 60
 20 tggcttaact gctagtaaga catttcagaa ctgagactgg tgtacgcatt tcatgcaagc 120
 cattaccact ttacctgaca ttttggacag agattagaaa tagtttcgta ctacctgcaa 180
 gttgcaactt gaaaagtga atttgttctt tgctaataata ttggcgtgta attcttttat 240
 gcgttagcgt aaaaagttga aatttgggtc aagttactgg tcagattaac cagtaactgg 300
 ttaaagttga aagatggtct tttagtaatg gagggagtag tacactatcc tcagctgatt 360
 taaatcttat tccgtcggtg gtgatttctt caatctccca acttagtttt tcaatatatt 420
 25 cataggatag agtgtgcata tgtgtgttta tagggatgag tctacgcgcc ttatgaacac 480
 ctacttttgt actgtatttg tcaatgaaaa gaaaatctta ccaatgctgc gatgctgaca 540
 ccaagaagag gcgatgaaaa gtgcaacgga tatcgtgcca cgtcggttgc caagtcagca 600
 cagacccaat gggcctttcc tacgtgtctc ggccacagcc agtcgtttac cgcacgttca 660
 catgggcagc aactcgcgtc atcttcccac gcaaaacgac agatctgcc tatctggtcc 720
 30 caccatcag tggccacac ctcccatgct gcattatttg cgaactccat cccgtcctcc 780
 acgcccacac accgcacac ggtcgcgata gccacgaccc aatcacacaa cgccacgtca 840
 ccatatgtta cgggcagcca tgcgcagaag atccgcgac gtgcgtgtcc cccgtgtcgg 900
 ttacgaaaaa atatccacc acgtgtcgtc ttacaggac aatatctcga aggaaaaaaa 960
 tcgtagcggg aaatccgagg cagcagctgc gattggctgg gaggcgtcca gcgtggtggg 1020
 gggcccaccc ccttatcctt agccgtggc gctcctcgtc cctcgggtcc gtgtataaat 1080
 35 accctccgga actcactctt gctggtcacc aacacgaagc aaaaggacac cagaaacata 1140
 gtacacttga gctcactcca aactcaaca ctcacacca 1179

 <210> 15
 40 <211> 1808
 <212> DNA
 <213> Oryza sativa

 <220>
 45 <221> misc_feature
 <223> PRO0133 - chitinase Cht-3

 <400> 15
 tttggcgcgg ggcagaagag tggactttaa ctttcttttt aataaaatct ccaattaata 60
 tgtaattata atatactttt aatcaaaaaca tgcaaagcta gcagtattta catcactaga 120
 50 agtaaatctt tcttgctcat gatgcttcag ccggacggaa ccctaaaata tagatggggc 180
 ggatacactc gattaaaaca gctaattgca acacatatca tataaggttt tgggaattcat 240
 accaaatgct ccgaaattcg tctatttcga tgaggcccaa gacatgacct cctgtttcgc 300
 ccatagttta tgggtgttgg taaaatttgg ttaaaatctg tctattttag taggtcccga 360
 aattcttatg caattgaatc ctagaaccct atcatattta tattgcaatt gcacaaaaat 420
 aatgtgcaat caatatattc caattgcaat acatatcaag catgaggtgt aatacatatc 480
 55 cagccgctag cactgggtct gttgaggtgc ttcttgacgc aacagctgca atctgttttg 540
 ctaggctgtt ggcgccaggc actgctgtcg tgctgcaaca atggcacatt cgtcgagcac 600

EP 1 801 223 A2

acaaccgccc ctatgcacag cgcaagctcg ctgccttggg ccgtgggtcc agtgttgcac 660
 caaggcttag tggattgagc gagaagacga actgacaatg ccaaagatgc gatgctgcga 720
 gtgtggactg cgaagatga atcgagatca atcaattcgt tatgcttgaa aggctggaat 780
 aactgatcag ttggctggat cgatgggtatg tactagataa tatgcggtct aggcctagac 840
 5 caagaagcag aagaggagtc gggcggggag tgtggggcga cgtaggctgt agctgggccc 900
 gccgccccag gccgcctaata gagtgtgtcc gcccttggcc tgacacgatg ggtaattaaa 960
 tagttatgca tgtccctctt tgtctaaaca atatgtataa aattgacgat atcttgggca 1020
 aaatcactgg gcatggcaca caggagagct actttagcga catgaatcta ggcgaaaatc 1080
 tattgaacca aaaatcgact gtaatctcat gaaaattttc gtcataatta tagcaaaatc 1140
 10 gttgttggat tgattgcacg agaaaacaga agaagggagc taggtgatata tatattgttt 1200
 tgttgccctac ataaatctta aagcaatcga atggtctaaa atttacaaga tttttaaaga 1260
 ggttttcgta ccgtatagac ccgcggccggg tcaaaacttat ttggctcgtcg ctgggttgttt 1320
 gtagcacgcc agctccatat atgtggattg cagctgggtct atgataagtt cggctcgatct 1380
 gagatcaatc tatcaatcgt caaccctttg cctttgttag cgagctagcg tgtacacatt 1440
 15 tcaattatat atggtgcatg catggcatcc acgcctccac ggtcaacgtg gaaatatctc 1500
 tggaacctta ctttttctaa ataactgaac ggattggagg caggagacaa atttgaccaa 1560
 cacaatatat ccacgacggc tagacaatac tagtagatgc atgcatggaa ggatattgta 1620
 gtacttggtta atcgtggaaa ctttggtaat gcgaatgcat ttcaattcgt tgcgtgaagat 1680
 cgatgcacca tgcatacca tctctatata aagccatgcg atcccaccga ttcttgacaa 1740
 20 cactactagct acttctactt ctatcatacc aaacaaacta gcttaatttg cattgcatca 1800
 cattgccg 1808

 <210> 16
 <211> 1828
 <212> DNA
 25 <213> Oryza sativa

 <220>
 <221> misc_feature
 <223> PRO0151 WS118

 30 <400> 16
 gcttgagtca tagggagaaa acaaatcgat catatttgac tcttttccct ccatctcctt 60
 taccggcaaa aaaagtagta ctggtttata tgtaaagttaa gattctttta ttatgtgaga 120
 tccggcttaa tgcttttctt ttgtcacata tactgcattg caacaattgc catatattca 180
 35 cttctgccat cccattatat agcaactcaa gaatggattg atatatcccc tattactaat 240
 ctagacatgt taaggctgag ttgggcagtc catcttccca acccaccacc ttcgtttttc 300
 ggcacacata ttttcaaaact actaaatggt gtgtttttta aaaatatatt caatacaaaa 360
 gttgtcttaa aaaattatat tgatccattt ttttaaaaaa aatagctaata acttaattaa 420
 tcacgtgta aaagaccgct ccgttttgcg tgcaggaggg ataggttcac atcctgcatt 480
 accgaacaca gcctaaatct tgtgtctag attcgtagta ctggatata taaatcatgt 540
 40 tctaagttac tatatactga gatgaataga ataagtaaaa ttagaccacac ctttaagcttt 600
 gatgaagtta ctactagctg cgtttgggag gacttcccaa aaaaaaaagt attagccatt 660
 agcacgtgat taattaagta ctagttaaaa aaacttaaaa aataaattaa tatgattctc 720
 ttaagtaact ctctataga aaacttttac aaaattacac cgtttaatag tttggaaaat 780
 atgtcagtaa aaaataagag agtagaagtt atgaaagtta gaaaaagaat tgttttagta 840
 45 gtatacagtt ataaactatt cctctgttc taaaacataa gggattatgg atggattcga 900
 catgtaccag taccatgaat cgaatccaga caagtttttt atgcatattt attctactat 960
 aatatacac atctgtcta aatatcttat atttcgaggt ggagactgtc gctatgtttt 1020
 tctgcccggt gctaagcaca cgccaccgcc gatcgggga cgctctggc cttcttgcca 1080
 cgataattga atggaacttc cacattcaga ttcgataggt gacgctcgac tccaagtgtc 1140
 ttgcacaaaa caactccggc ctcccgccca ccagtcacac gactcacggc actaccacc 1200
 50 ctgactccct gaggcggacc tgccactggt ctgcatgcga agctatctaa aattctgaag 1260
 caaagaaagc acagcacatg ctccgggaca cgcgccacc ggcggaagag ggctcgggtg 1320
 ggcatctca cagccgcata tcgatttca caagccgcc atctccaccg gcttcacgag 1380
 gctcatcgcg gcacgaccgc gcacggaacg cacgcggccg accgcgcgc ctcatgccc 1440
 gagcccatcc gccgcgtcct cctttgtcct ttgccgtat cctctcgtc gtatcccgtt 1500
 55 tctctgtctt ttgtccccc gcgcgcgcca gttcggagta ccagcgaac ccggacacct 1560
 ggtacacctc cgccggccac aacgcgtgtc cccctacgt ggccgcgcag cacatgccc 1620

EP 1 801 223 A2

tgcgcgacac gtgcacctcc tcataccaaac tctcaagtct caacgggtcct ataaatgcac 1680
 ggatagcctc aagctgctcg tcacaaggca agaggcaaga ggcaagagca tccgtattaa 1740
 ccagcctttt gagacttgag agtgtgtgtg actcgatcca gcgtagtttc agttcgtgtg 1800
 ttggtgagtg attccagcca agtttgcg 1828
 5
 <210> 17
 <211> 1267
 <212> DNA
 <213> Oryza sativa
 10
 <220>
 <221> misc_feature
 <223> PRO0169 - aquaporine
 15
 <400> 17
 cgtcctcctt ttgtaacggc tcgcaaatac aatgggttgt ttagattcat gtcatttttaa 60
 atcatattat tttttataaa gttatcaaaa tgtacatata tttatttatt tttaccaaac 120
 tttactaaat gagataatcc aacaaatggc atttaaagcg ttcaaatacca agaaatgcc 180
 tcgcggttat gcttccgtcc gtttcacggc gttaaaatac aatgttcac cttataacact 240
 taatgggtgtg gaatggacgg aaccctaacg gcgatggcat ttttgggata aagtcggttg 300
 20 tacgatggca tttcttagaa ctcatatttg tcgatggcat tttttgaatt tggatgattg 360
 tcaatgggtat tttttggatt atctcttagt aaatacataa ggaatcatgc caaaacttga 420
 caatattgtc aacttatcaa aatttaattg ggattatttt ggcgataata tgaacagccc 480
 ttacattttc gaagaattat agctcaataa tggctatggc cctgttttga ttccggagggc 540
 tattttaatag cctccggaa tcttgctatt taagagtatt aaacgtagat tactgataaa 600
 25 actcattcca taacccctac gctattctac gagacgaatc taacgaggtt tattaatcca 660
 tgatttgcta cagtaatcag ccgctaactg tggattaata tacatcatta gattcgtctc 720
 gtaaaatagg ctagggatta tggaaatcgg tttatcggta atctatgttt aatacttcta 780
 aatagcaaga ttccgaaggg ctattttaata gctcggagca tccaaacaag gcctatgttt 840
 agatccaaac ttccaaactt ttctatcaca ttaaaactgtc atacatacat aacttttcag 900
 30 tcacatcgta ccaatttcaa cccaaacttt caacttttga agaactaaac acagcatatg 960
 acagtgcagt tcagctcaat tttgttcgga gcctaaaaaa aagaaaagaa aaaaagctca 1020
 atttggataa ggctatgaat aaactcaaaa aagcatccaa cctaaccacc aactggccc 1080
 accagggccc acgctccact cccgtgatca tcacctcctt ccttttccag aaccaccttc 1140
 tctctccttc ctctcttctc tcttcagtgt actctgcctt tataacaccc tactcctctc 1200
 35 tctcacctcc accatctagc tcactcacac agtctccact cacacgcatt gcagaggaga 1260
 ggcgaca 1267
 <210> 18
 <211> 1130
 <212> DNA
 40 <213> Oryza sativa
 <220>
 <221> misc_feature
 <223> PRO0170 - High mobility group protein
 45
 <400> 18
 catgcggtta atgtagatgc tcactgcgct agtagtaagg tactccagta cattatggaa 60
 tatacaaagc tgtaatactc gtatcagcaa gagagaggca cacaagttgt agcagtagca 120
 caggattaga aaaacgggac gacaaatagt aatggaaaaa caaaaaaaaa caaggaaaca 180
 catggcaata taaatggaga aatcacaga ggaacagaat ccgggcaata cgctgcgaaa 240
 50 gtactcgtac gtaaaaaaaaa gaggcgcatt catgtgtgga cagcgtgcag cagaagcagg 300
 gatttgaaac cactcaaate caccactgca aaccttcaaa cgaggccatg gtttgaagca 360
 tagaaagcac aggtaagaag cacaacgccc tcgctctcca cctcccacc caatcgcgac 420
 gcacctcggc gatcggtgac gtggcctcgc ccccaaaaaa tatcccgcgg cgtgaagctg 480
 acaccccggg cccacccacc tgtcaagttg gcacatgttg gttatgggtc ccggccgcac 540
 55 caaaatatca acgcggcgcg gcccaaaatt tccaaaatcc cgcccaagcc cctggcgcg 600
 gccgctcttc caccaggtc cctctcgtaa tccataatgg cgtgtgtacc ctcggtgtg 660

EP 1 801 223 A2

tgtacgtggg cggggttacc tgggggtgtg ggtggatgac ggggtggccc ggaggaggtc 720
 cggccccgcg cgtcatcgcg gggcggggtg tagcgggtgc gaaaaggagg cgatcgggtac 780
 gaaaattcaa attaggaggt ggggggcggg gcccttggag aataagcgga atcgagata 840
 tgcacctgac ttggtttggc tctctctctt cttatccctt gtcctcgcaa ccccgcttcc 900
 5 tctctctctc tctctctctc tctctctctc tgggtgtgtg ggtgtgtccc tgcctccctc 960
 ctctctctc ctctctctt cctctctctc ttccccctc tcacaagaga gagagcgcca 1020
 gactctcccc aggtgaggtg agaccagtct ttttgcctga ttcgacgcgc ctttcacgcc 1080
 gcctcgcgcg gatctgaccg cttccctcgc cttctcgcga ggattcagcc 1130

 10 <210> 19
 <211> 1230
 <212> DNA
 <213> Oryza sativa

 15 <220>
 <221> misc_feature
 <223> PRO0171 - reversibly glycosylated protein RGP1

 <400> 19
 20 tagtaaccatt cttccctcgt gagcataaat gtattcatac aaaatagtaa aatgtatcct 60
 cacaaagatt gtaagtatat ctcgcaacta taaatatgtt gtcatttttag taacaattgt 120
 tcataaaata gtaatcatgt tctccataac agtaaattgac gaggcgttaa tagtggttta 180
 ggttctcatg attgtaaatg ttgagtcgct tgtagcggct taagatatag tagagagtat 240
 atctagtttt atcaagacaa acattgcgta atgcctcgga cctaataata aagtaggaat 300
 tttaaccttt gagaaactgt aaccaattga aactgcaagc tttaaaaaaa catctattgg 360
 25 aagtgatatt atatagacaa aataagtttc ttactcttac tctctcagtt tcaagttata 420
 aaatgttttg gctttgggtc aaatcaaaact tcttcaagtt taatcaagtt tatagaaaaa 480
 atagtaatat ccaagataaaa tttattataa aaatatattt aattattatt ttaataaaac 540
 taatttggtg atgtaaatat tactataattt gtctataaac ttagtcaaat ttaaaacagt 600
 ttaactttga ccaaagtcac aacatcttat aacctgaaat ggatggagta tttggttggt 660
 30 tctatttttag gaaacggcgc tttctttcca ttgattttga gataagcaga gctttaaacc 720
 actgccacta ttgtgcattt catttgattt aacactttta ccccttatct ccaataaaaa 780
 cgatattaag atacccttat cttttatcca ccgcttggaa caaaccaaaa aaaataaaaa 840
 ttcaaacctt ctacactggg acacacgttc tctctttcca tgcaccgaca ggtctctccc 900
 agatccaacc caaaataaat ttggacgcac cccaaaattc ggcaaacata tgacgcaaac 960
 caaaacaaaa taggcacaaa ataataaat cctcttatct aattaattat acacaatttt 1020
 35 ttttaaaaaa aaagcaaggc aagcgaagca aagcaaagaa ggaaacgaat aacaaagtcg 1080
 tgcctctccc ggagctcccg ctctataaat cgctcctcct cccacccac ccaaaccac 1140
 acacacctca cactcacca ccatcacctc ctctctctcc tctcttctct ccgcgcgcgc 1200
 gagatccagg gagagggaga gggagagatc 1230

 40 <210> 20
 <211> 1234
 <212> DNA
 <213> Oryza sativa

 45 <220>
 <221> misc_feature
 <223> PRO0173 - cytosolic MDH

 <400> 20
 50 gtttggttgg tgaccgcaat ttgctatacc aaaatcttag acacagttga attagctac 60
 actttatttag cacattggcc cgtgcgttat attgtcattt tctagccaaa gtttgccata 120
 attgtggcta acaaattgtt ggccacattt tggctacggt cgataggaca tgttcccaac 180
 ttctccttct cgtttttcgc gcgtacgctt tttcaaactg ttaaacggcg tgtttttcgc 240
 aaaatatttt ttacgaaag ttgcttaaaa aattatatta atctattttt tttaaaaaaa 300
 gtagctaaaa ctttaattaat ctcacgctag acgctgcttc gttttacgtg tcgggtaccc 360
 55 aacctcact cccgaacaca gcctttgtgt ggtttactac agttatagta aagctagtct 420
 ccatccaaac aatcctttag tccatataac ttgctatact ccaaaattcc actcgttcta 480

EP 1 801 223 A2

```

cggacatcac taatacgaag atcaagtgga agatagatat ttttaatgac atgttatttt 540
cagtgaacac ttgaggtcct cagatccac aaacacacat tttcgtagat aagttctgaa 600
atactccata cggcggttgt cagatgtca tgatcgtcgt tacccaagga agaagaaaag 660
agtggcatct tctccacgcc agtggtccca acggagcatc ttttcttccc ccacacggca 720
5   tcgacgtcac actttctggt gcaaacttta ataattagtc caaaaacaaa aaaagaattt 780
cggccacatc ttctcccgaa acgccagggt ggccccacct gcatcactga cagcctgtcc 840
ccacaacgcg cagtcgtgtc ccacactgtc aggatgttag cgtctccgtt gcaggtttcc 900
cagatcccat cgccgatctg tgggccagcg cccacgggtg cagccccgcy cacacctggc 960
tccaacccac ccaccccacg cgctccgtgg ccgacagcgt ggacccacct aggtggggcc 1020
10  caccgtcagt gggagatggg taggggagcg cccacgtggg agcaacgggg gttctccggg 1080
ctccccgtcg ccgagagggt aaataacggc caccgtttc cccctctctc gcaaaactca 1140
cccaaaagag cagcgtcgcc tctctctctc ccccttaacc cctacgcttc cagaaccttc 1200
togaagctcc cgctccccc ccccttccgc tcca 1234

<210> 21
<211> 1553
<212> DNA
<213> Oryza sativa

<220>
20 <221> misc_feature
<223> PRO075 RAB21

<400> 21
gtcaccaccg tcatgtacga ggctgcttca ccactgcctc actgccacca gcgtctcccg 60
25 ccgcgtgcaa tacaagaaga aacatcgaac ggtcatataa ggtaagacc actaccgatt 120
taacctatca tccccacaat ctaatccact tatttctctt cccatgatct tatcctctca 180
tttctcttca ctacttttgc atttgtagga aacacaatga caccgtcgaa gaaagctggg 240
ggagcaccgt agccgaat caccaaaaca cagaggggag gaggtcggca gcggccatgc 300
ggacggcgat gagacaacgc gacgcaaaga gggaggagga cgttggcgat catgctggtg 360
30 ttggcggagg aggtcactgg ccattgcgaat gacagcgggg cagcgcaaca caaaaagggg 420
ggaggatgcc ggcgaccacg ctagtaccat gaagcaagat gatgtgaaag ggaggaccgg 480
acgaggggtg gacctctgcc gccgacgtga agagcgtgat gtgtagaagg agatgttaga 540
ccagatgccg acgcaactta gccctgcaag tcaccgcact gcatatcgtt gcttgccctc 600
gtcctcatgt acacaatcag cttgcttacc tctccatact tgtcgtttgt tcccgtggc 660
cgaaatagaa gaagacagag gtgggttttg ttggagagtt ttagtggtat tgtaggccta 720
35 tttgtaattt tgttgacttt tattgtatta atcaataaag gtgtttcatt ctattttgac 780
tcaatgttga atccattgat ctcttggtgt tgcactcagt atgttagaat attcattccg 840
ttgaaacaat cttggttaag ggttggaaca tttttatctg ttoggtgaaa catccgtaat 900
attttgcgtt aaacaatttt tatccgacag caccgtccaa caatttacac caatttggac 960
gtgtgataca tagcagtcct caagtgaaac tgaccaccag ttgaaaggta taaaagtga 1020
40 acttattcat ctaaaagacc gcagagatgg gccgtggccg tggctgcgaa acgacagcgt 1080
tcaggcccat gagccattta ttttttaaaa aaatatttca acaaaaaaga gaacggataa 1140
aatccatcga aaaaaaaaaa ctttccctacg catcctctcc tatctccatc cagggcgagc 1200
actcatccaa accgtccatc cagcgcaca gtacacacac atagttatcg tctctcccc 1260
cgatgagtcg ccaccctgtt ctccgagaaa cgcctcgccc gacaccgtac gtgcgccacc 1320
45 gccgcgectg ccgcctggac acgtccggct cctctccgcg cgcgtggcc accgtccacc 1380
ggctcccgca cagctctccc tgtctccctc caccatgcc gtggcaatcg agtcatctc 1440
ctcgccctct ccgcttata aatggcgccc accaccttca cctgcttgca caccacagca 1500
agagctaagt gagctagcca ctgatcagaa gaacacctcg atctctgaga gtg 1553

<210> 22
50 <211> 1087
<212> DNA
<213> Oryza sativa

<220>
55 <221> misc_feature
<223> PRO077 - Cdc2-1

```

EP 1 801 223 A2

```

<400> 22
cagacaccta gaatatagac attcccaaaa aataatcact atgcatcagc atcactatac 60
atgacttggg tctagtgatg gaagtggata gttccactac ctacataaaa acccactact 120
5 agtttattac ttttcacatg atagcataaa atttaaagaa aaaataaaca gaagtggaat 180
aagcgaaaaa ccccgcttac ccgccccatt tacatcccta cttggatcct gcatgtcagt 240
aagatatcag aattatatgt tttagaatta tatgtttttt tgggaaggtgg aaatcggatt 300
attagacgca acataccaag tggcgtatac ttggcttcac tctttccatc agagcaagcg 360
taaaagatca cgtattcacg tcacatggag taactgagcg aatttttttc atttttaaat 420
10 ttttgttttt taatatattac ataaaatatta taccggcgaa aatattttaca aaagtagacc 480
ctgctgccct tctccttctc gagaagagcg gcaggggtgat gtcagggaca gaaataaact 540
ccaaaaatgc atttttggct gggcgaaaat tgcacttacc cccttgctgc cctctacaaa 600
ggttgcaagg gacctcagtg caaaatacgc acaccttgcc gtcctccact tggacggcat 660
gggctatttc tgtaaatatt ttggatggta taatatctct gtaaatatta aaaaataaaaa 720
15 atttaaaaat gaaaaaatlc tatctgggct cccttctctc atctcacacg gccaccaca 780
caatcccggc ccacatatct cctgggcccc ttcccggtg aatggagacg gccattggc 840
gcgcacatgc gcaaaagcgt acacacgatt cgaaaattga aatctcaaaa agcggccgtt 900
agagcgcgtc ccctccaacg gctatcccca atacaaaaga tcaactcgaat cccccccaaa 960
tcgaccaaac cctaaatcca cgcgcattcc acaccacca accagcgaga gagagatggc 1020
20 ggcgtccac caccaggcgg cggcggcgcg ggtgacgacg acgacggacg ggggcgagct 1080
gcgggcg 1087

<210> 23
<211> 1272
<212> DNA
25 <213> Oryza sativa

<220>
<221> misc_feature
<223> TC89946 (PRO0110)
30

<220>
<221> misc_feature
<222> (17)..(17)
<223> n = any nucleotide
35

<220>
<221> misc_feature
<222> (50)..(50)
<223> n = any nucleotide

40 <400> 23
tttgacgact gaatcgnngc tcgcctctgc ggcggcgcgct ctagattagn gtttcccttg 60
tctgttgtaa ttccggcacga gggctgatca agagctctta attagctagc tagtgattag 120
ctgcgcttgt gatcgatcga tctcgggtac gtagcaatgg cgtccaaggc gttcgtctctg 180
ttcctggcgg tgaacctcgt cgtgctcggg gtggcaagcg cctgcggcgg cagcccgctc 240
45 tgcccgacgc cgacgccgtc gaccccgaca ccgtcaacgc cgacgccgac gccgtcggcg 300
ttcgggaggt gccccgcgca cgcgctgaag ctgggcgtgt gcgccaacgt gctggggcctg 360
atcaaggcca aggtgggcgt gcctccggcg gagccgtgct gcccgctgct ggaggggctc 420
gtcgacctcg aggcggcggt gtgcctctgc acggccatca ggggcaacat cctcggaatc 480
aacctcaacc tccccatcga cctcagcctc atcctcaact actgcgga gaccgtcccc 540
accggcttca agtgctaagc agcgtgcata tgcaatgcct gcatgggttg atcctacgta 600
50 cggtgattag ttggctttga cgactcttga tttgatttgc ttgctgctct gtttatttgc 660
tactacgtta cgtacgtact ttgcatgcaa cgcaacgcat gatcgatcgt gcatgctggc 720
tgtttgtacg tatcacggta ccagtttggg ttctctctgt actctctcct ttgtcttctt 780
tgtagtactc ttattccgcg tatccgtacg tgcgcatctt ttgtaaggcg cgggtgtagc 840
ttgtgtgccg gtaaccaact ctaataaagc tatgggtgga acttcaaaaa aaataaaaaa 900
55 aaaactggag gggggggccc ggtccaattt agactataat gagtttaaca ccccgctcat 960
cgccgaaga taacaacacc gggcttggaa aacctagact gcccaactaa tggacggaag 1020

```

EP 1 801 223 A2

acagactcct ggactgaaac tgaacgaaac aagaccaccc accccatcta accacagcca 1080
 cctaccgcca aagattccaa taatgtgaat cagtcggtaa tagaacactc ctcttgtagc 1140
 attttactgc ccgcgccacc cctcggtagc cacttatata tatcggggcg tagtaatttc 1200
 ctggttcggt cactccctc atcgcacctg ctagtctgtg cttacatacg tgcgtcctct 1260
 5 tattatcgag cg 1272

<210> 24
 <211> 2425
 <212> DNA
 10 <213> Oryza sativa

<220>
 <221> misc_feature
 <223> TC90358 (PRO0005)

15 <220>
 <221> misc_feature
 <222> (1558)..(1558)
 <223> n = any nucleotide

20 <400> 24
 cccacattga ataattatit taaataatit aagttttitit tttttggctt tagatatatt 60
 cccaatcccc aacctcccaa taatccgatc tctcccagtt ctgttcggat caaggctgtg 120
 tcgatcgcaa aaaagaaaaa aaaaacaatt tccttttggg gtggttcacg tgttgatcac 180
 ttctttgttt ccgcgctttt gttggggatt cgattttcgg gtttaagatt tctacacgat 240
 25 ggccttgaac ttggctcaga gcgcgcgcgc ggcagcgtgc ttcgcgaccg ccggtgatgc 300
 gcggcgagct gcttcggtgg tcgccatgcc gtcgtcgtcg tcgtcggcca cgacgagcct 360
 gaggatgaag aggcaggcgg cgtgcgagcc ggtggcgtgc cgggcggtgg ccaggcacgt 420
 ggccggcgcg gcggcgagca gcaggaggaa cggcgtgcgg gtgttcgtga tgatgccgct 480
 ggacacggtg agcaagtgcg ggagcgcgt gaaccggagg aaggcggtgg cggcgagcct 540
 gcggcgctg aagagcgccg gcgtggaggg gatcatggtg gacgtgtggt ggggcacgt 600
 30 ggagagcgag ggccccggcc ggtacaactt cgacggctac gtggagctca tggagatggc 660
 ccgcaagacc ggcctcaagg tccaggccgt catgtccttc caccagtgcg gcggcaacgt 720
 cggcgactcc gtcaacatcc cgtcccgag gtgggtggtg gaggagatgg agaaggacaa 780
 cgacctcgcc tacaccgacc aatggggacg ccgcaacttc gagtacatct cctcggctg 840
 cgacgccatg cccgtcttca agggccgcac gccgctcgag tgctacacgg acttcatgcg 900
 35 cgccttcgcg gaccacttgc cctccttcct cggcgacacc atcgtcgaaa tccaagtgcg 960
 catgggcccc gccggcgagc ttcgggtacc gtcctaccg gagagcaacg gcacctggag 1020
 gttccccggc atcggcgccct tccaatgcaa cgacaggtac atgcgtagca gcctgaaggc 1080
 ggccggcgag gcgaggggca agccggtagt ggggccacgg cgggcccagc gacgccggcg 1140
 gctacaacaa ctggccggaa gacacggtgt tcttcgcggc cgactgcggc ggggtggagca 1200
 40 ccgagtacgg cgagttcttc ctgctgtggt attcgcagat gctgctggag cacggcgagc 1260
 gcgtgctgtc gggcgcgacg tccgtgttcg gcgacggcgc cggcgccaag atctcgttca 1320
 aggtggccgg catccactgg cactacggca cgcggtcgca cgcgcggag ctcacggcgg 1380
 ggtactacaa cagcggcgac cgcgagcggc tacctcccga tcgcgcgcat gctggcgcg 1440
 caccggcgccg tgcctcaactt cactgcgtg gagatgcgcg accacgagca gccgcaggag 1500
 45 gcgcagtgca tgcccagggc gctcgtcagg caggtggccg ccgcggcgcg cgcggcgnga 1560
 cgtcgggctc gccggggaga acgcgctgcc gcggtacgac ggcacggcgc acgaccaggt 1620
 ggtcgcgcgc gccgcgacc gcgcggcgaa ggaccggatg gtgccttca cctacctccg 1680
 gatggggccc gacctcttcc acccgacaa ctggcgccg ttcgtcgcct tcgtccgcg 1740
 catgtccgag tccggctcgc cgcgggaggc cgcgagagc gccgcgcac gcgtcgcgca 1800
 ggccaccggc tcgctcgtgc acgagggcgc ggtcgcgctc cggagctagc accggtcaga 1860
 50 cgtcatata caccgtcgcc tcgaggtcgg attccgatgt gggatcattc gatctccctt 1920
 ttttttttct tctttttgcc attttgtaca gccttttggg gagctttgga tttgtgcttt 1980
 ttgtctcggg agggaaaaccg ctctggaggt cgaagagagc gtcattttcc tccggttgaa 2040
 gatcacgaat catttacgtt agagatgatg taattaagca gggaggggag gggaacacac 2100
 acacactggc actcaaaagt tgttgtcacg cttggggaat atatccattt ccagccaaaa 2160
 55 aaaaaacgca gaaatgcgtt gtgttcttgc gctctggttc gttgctgctg tgggtcagat 2220
 tcagctggtg aaaaaactac agtactactg aaactgaaac tactagagcc tagagggaga 2280